

# Bilgisayar Ortamında Biçimsel Ontoloji Oluşturulması

## Construction of Formal Ontology on Computer

Yelda FIRAT, Yılmaz KILIÇASLAN, Özlem UÇAR

### ÖZ

Genel bir bakış açısıyla, biçimsel bir dil kullanarak tanımlanmış gerçeklik modellerine biçimsel ontoloji denir. Bu genel amaçlı modeller, özel uygulama alanlarında da kullanılabilirler. Ontolojiler doğaları gereği büyük olmaları nedeniyle, insan emeği ile oluşturulmaları halinde hata ve eksiklikler barındırma potansiyeli taşırlar. Ayrıca, bu şekilde oluşturulmuş olan bir ontolojinin, belli bir uygulama alanına uyarlanması da güç olacaktır. Bu çalışmada, insan emeği kullanımından kaynaklanabilecek olası sorunlardan korunmuş bir bilgisayarlı biçimsel ontoloji oluşturma uygulaması tanıtılacaktır. Uygulamada, biçimsel kavram analizi çerçevesinde derlem tabanlı işaretleme yöntemiyle özel uygulama alanı olarak Türkçenin anlamsal çözümlenmesi hedeflenmiştir. Bu tip doğal dil işleme uygulamalarının yanında biçimsel ontolojiler, bilgiyi anlamsal yapılar şeklinde açık ve anlamlı bir şekilde sunmalarından dolayı eğitim alanında da çok fazla bir şekilde kullanılırlar. Bunun en önemli nedenleri; yapılandırmacı yaklaşımın yayılışı ve öğretim amacıyla kullanılan kavram haritaları gibi bilgi teknolojilerinin kullanımınıdır.

**Anahtar Sözcükler:** Biçimsel ontoloji, Biçimsel kavram analizi, Bilgi teknolojileri, Anlamsal yapılar

### ABSTRACT

From a general point of view, the defined reality models by using a formal language are called formal ontology. These general purposeful models can be also used in the special fields of application. Because ontologies are large-scale by their nature, they have the potential of hosting errors and deficiencies in case they are generated by human labor. Besides, an ontology formed such a way will be difficult to be adapted to a application specific. In this study, an application of a computerized formal ontology formation saved from probable problems which are likely to arise using human labor will be introduced. In the application, semantic analysis of Turkish has been targetted as a special application field using corpus-based marking method in the framework of formal concept analysis. Beside the applications of these kinds of natural language processing formal ontologies are used considerably in the field of education as they present the information in the way of semantic structures clearly and meaningfully. The most important reasons of this are the use of the information technologies such as expansion of constructivist approach and the concept maps used for educational reasons.

**Keywords:** Formal ontology, Formal concept analysis, Information technologies, Semantic structures

### GİRİŞ

En geniş anlamıyla dil, insanların aralarında anlaşmalarını sağlayan, onların duygu ve düşüncelerini doğrudan ya da dolaylı olarak bildirmeye yarayan işaretler bütünüdür. Genellikle, dilin üç bileşenden oluştuğu iddia edilir: Dilin öğeleri arasındaki ilişkileri kapsayan biçimsel/dizimsel bileşen, dilin dış dünya ile

ilişisini barındıran anlamsal bileşen ve ifadeleri konuşmacı ve dinleyici bağlamına yerleştiren pragmatik bileşen.

Dilin özellikle anlamsal çözümlenmesi, gerçekliğin bir modelini, yani bir ontoloji kullanmayı gerektirir. Bilgisayarlı dilbilim ya da doğal dil işleme uygulamaları ise, bu ontolojinin biçimsel olmasını zorunlu kılar. Söz konusu ontolojik modeller çok büyük

#### Yelda FIRAT (✉)

Çanakkale Onsekiz Mart Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Çanakkale, Türkiye  
Çanakkale Onsekiz Mart University, Vocational School of Technical Sciences, Çanakkale, Turkey  
yelda1367@gmail.com

#### Yılmaz KILIÇASLAN

Trakya Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Edirne, Türkiye  
Trakya University, Faculty of Engineering and Architecture, Edirne, Turkey

#### Özlem UÇAR

Trakya Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Edirne, Türkiye  
Trakya University, Faculty of Engineering and Architecture, Edirne, Turkey

**Geliş Tarihi/Received :** 23.10.2012

**Kabul Tarihi/Accepted :** 07.12.2012

olmaları nedeniyle, insan emeğiyle oluşturulmaları halinde, birçok hata ve eksikliği barındırma riski taşırlar. Ayrıca, bu şekilde oluşturulmuş bir ontolojinin uygulama gereksinimlerinden doğabilecek uyarlamaları da güç olacaktır.

Bu çalışmanın amacı, bir bilgisayarlı biçimsel ontoloji oluşturma uygulamasını tanıtmaktır. Uygulama, biçimsel kavram analizi çerçevesinde derlem tabanlı işaretleme yöntemiyle geliştirilmiştir. Özel uygulama alanı olarak Türkçe'nin anlamsal çözümlenmesi hedeflenmiştir.

Makalenin izleyen bölümünde, sunulan çalışmaya dair artalan bilgileri sunulacaktır. Özellikle, özel uygulama alanı olarak seçilmiş cümle ögelerinin tematik anlam analizi ile ilgili açıklama ve örnekler bu bölümde yer alacaktır. Sonrasında, tematik rol analizinin kavram latisleri ile nasıl modellenebileceği tartışılacaktır. Bu kuramsal temel üzerinde gerçekleştirilmiş derlem tabanlı bir biçimsel ontoloji oluşturma uygulaması izleyen bölümün konusu olacaktır. Sonuç bölümünde, makalenin özeti ve gelecekte yapılabilecek çalışma önerileri sunulacaktır.

### Tematik Roller

Birçok dilbilim kuramında, 'tematik rol' terimi bir ad öbeğinin tümcenin eylemi ile girdiği anlamsal ilişkiyi ifade etmek için kullanılır. Aşağıda literatürde kullanılan belli başlı tematik rollerin bir listesi yer almaktadır (Jackendoff, 1972). Altı çizili ad öbekleri tanımlanan tematik rolü taşımaktadır:

TEMA (THEME) – Hareket halinde bulunan veya bir yerde olan

(1) a. Kedi odada.

b. Top yuvarlandı.

YAPAN (AGENT) – Bir eylemi başlatan, yapan

(2) Kedi camı kırdı.

ETKİLENEN (PATIENT) – Bir eylemden etkilenen

(3) Kedi camı kırdı.

KAYNAK (SOURCE) – Hareketin çıkış noktası

(4) Can İstanbul'dan Ankara'ya gitti.

HEDEF (GOAL) – Hareketin yöneldiği şey

(4) Can İstanbul'dan Ankara'ya gitti.

KONUM (LOCATION) – Bir şeyin bulunduğu yer

(5) Kedi odada.

DENEYİMLEYEN (EXPERIENCER) – Bir olayı hisseden, algılayan veya fark eden

(6) a. Can'ın başı ağrıdı.

b. Kedi topu gördü.

ARAÇ (INSTRUMENT) – Bir eylemi gerçekleştirmek için kullanılan

(7) Can ekmeği bıçakla kesti.

Birçok dilbilimci, tematik rolleri ve bu rollerin biçimsel/dizimsel özelliklerini gramer teorisinin önemli bir parçası olarak tanımlarlar (Gruber, 1965; Fillmore, 1976; Dowty, 1991; Jackendoff, 1993).<sup>1</sup> Özellikle ögelerin sınıflandırılmasında, tematik rollerin önemi çok eski yıllara, Panini'nin Karaka teorisine kadar uzanmaktadır (Dahiya, 1995; Kasper, 2008). Doğal dil işleme uygulamalarında da, verilen bir cümlenin anlamını çözümlenmede ögelerin tematik rollerini belirlemek oldukça önemlidir (Gildea & Jurafsky, 2002). Bugüne kadar yapılan çalışmalarda, soyut tematik roller üzerinden dilin ögelerinin anlamsal rollerine ulaşıldığı görülmektedir.

### Tematik Rollerin Kavram Latisleriyle Modellenmesi

Biçimsel kavram analizini kullanmak suretiyle elde edilecek kavram latislerini (Galois Latisleri) kullanarak, matematiksel bir tematik rol hiyerarşisi oluşturulabilir (Kılıçaslan & Güner, 2011). Biçimsel kavram analizi ilk kez Ganter ve Wille (1996) tarafından geliştirilmiştir. Biçimsel bağlamda verilen veri, iki kümeden ve bu iki kümenin ilişkisini gösteren ifadeden oluşur (Kılıçaslan & Güner, 2011). Kümelerden birisi özellikleri barındırırken, diğeri bu özellikleri taşıyan nesnelere içerir. Bu şekilde elde edilen kavram latisleri, hem tematik rolleri hiyerarşik bir organizasyon içinde modellememize imkân verirken, hem de bir ad öbeğinin birden fazla tematik rol taşımasına da izin verir. Bir örnek ile açıklayalım:

(8) Öğretmen öğrenci-ye bir kitap verdi.

cümlesinde, 'öğretmen' YAPAN ve KAYNAK, 'bir kitap' TEMA, 'öğrenciye' sözcüğü ise hem HEDEF hem de ETKİLENEN rolünü üstlenir. Yani, verilen cümlede 'öğrenciye' ve 'öğretmen' iki tematik rol almıştır. Şekil 1'de gösterildiği gibi, bu cümleye ait ögelerin her birinin tematik rol kümesi bir düğüm ile temsil edilir.

Şekil 1'deki nesnelere kümesini oluşturan harfleri açık bir şekilde ifade etmek gerekirse; ö(öğretmen), öğ (öğrenciye), k (kitap) dır.

### Uygulama

Bu bölümün amacı, bir bilgisayarlı biçimsel ontoloji oluşturma uygulamasını tanıtmaktır. Uygulama, biçimsel kavram analizi çerçevesinde derlem tabanlı işaretleme yöntemiyle geliştirilmiştir. Kullanılan derlem 1033 cümleden oluşmaktadır. Özel uygulama alanı olarak Türkçe'nin anlamsal çözümlenmesi hedeflenmiştir.

İşaretleme ise, verilen cümlelerin parçalara ayrılarak, her bir parçanın veya ögenin anlamsal ve biçimsel/dizimsel özelliklerinin örneklenmesidir. Kendi derlemimizden seçilen (9) nolu cümleye ait işaretleme tablo 1'de görülmektedir.

(9) Şölen başkan-ı, şarab-a dönüş-müş su-yu tat-tı.

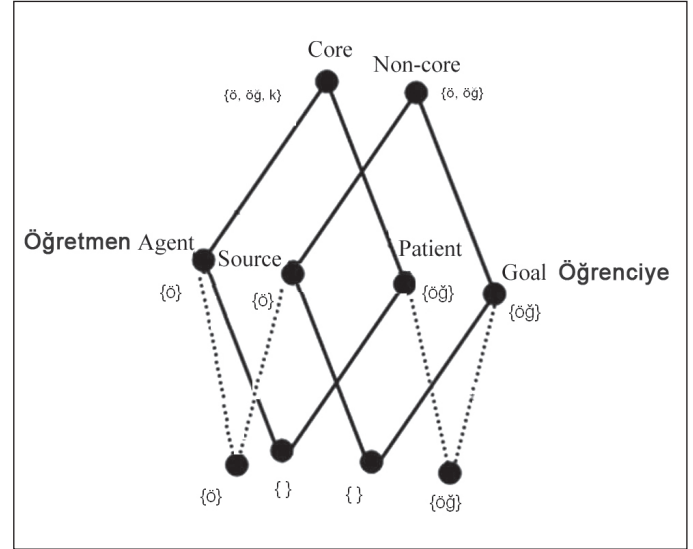
<sup>1</sup>Tematik rollere ilişkin literatürün kapsamlı bir tartışması Kasper (2008)'de bulunabilir.

**Tablo 1:** Örnek İşaretleme

[Theme tematik rolü]	=>	'suyu'
[Agent tematik rolü]	=>	'Şölen başkanı'
[Goal tematik rolü]	=>	'şaraba'
[T]	=>	'tı'
[VP]	=>	'tat'
[Case]	=>	'yu'
[NP]	=>	'su'
[ASP]	=>	'müş'
[VP]	=>	'dönüş'
[Case]	=>	'a'
[NP]	=>	'şarab'
[Case]	=>	'ı'
[NP]	=>	'başkanı'
[NP]	=>	'Şölen'

Bu tabloda da görüleceği üzere (9) nolu cümle için öğelerin tematik rol ve biçimsel/dizimsel işaretlemeleri yapılmıştır. Bu cümle için biçimsel/dizimsel özellikleri ise şunlardır:

- ✓ İsim sözcük öbeği (Noun Phrase-NP)
- ✓ Fiil sözcük öbeği (Verb Phrase-VP)
- ✓ Hal eki (Case)
- ✓ Zaman eki (Tense-T)
- ✓ Görünüş eki (Aspect-ASP)

**Şekil 1:** Latislerde tematik rollerin gösterimi.

Bu tip işaretlemeler için bilgisayar ortamında hazırlanmış olduğumuz işaretleme ara yüzü ("annotation tool") PHP dili ve MySQL'de oluşturulan veri tabanı kullanarak gerçekleştirilmiştir. Bu ara yüz web ortamında çalışan bir sayfa konumundadır.<sup>2</sup>

Şekil 2'de gösterilen ara yüz ile elimizdeki derleme ait Türkçe cümleler parçalara ayrılarak tematik rolleri işaretlenmiş, biçimsel/dizimsel analizleri yapılmıştır. Bu işaretlemelerden elde edilen bilgilerle oluşturulan veri tabanı, cümlelere ait öğelerin

**İsa ona şu karşılığı verdi:**

**Selected characters :**

**Selected Positions :**

[Verb]	=>	"ver"
[Agent]	=>	"İsa"
[Goal]	=>	"ona"
[Theme]	=>	"şu karşılığı"
[T]	=>	"di"
[VP]	=>	"ver"
[Case]	=>	"ı"
[NP]	=>	"karşılığı"
[AdjP]	=>	"şu"
[Case]	=>	"na"
[DP]	=>	"o"
[NP]	=>	"İsa"
[Subj]	=>	"İsa"
[Discourse]	=>	"İsa ona şu karşılığı verdi:"
[Active]	=>	"İsa ona şu karşılığı verdi:"
[Ext]	=>	"İsa"
[Comp]	=>	"ona"
[Obj]	=>	"şu karşılığı"
[Patient]	=>	"ona"
[Source]	=>	"İsa"

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

Delete label

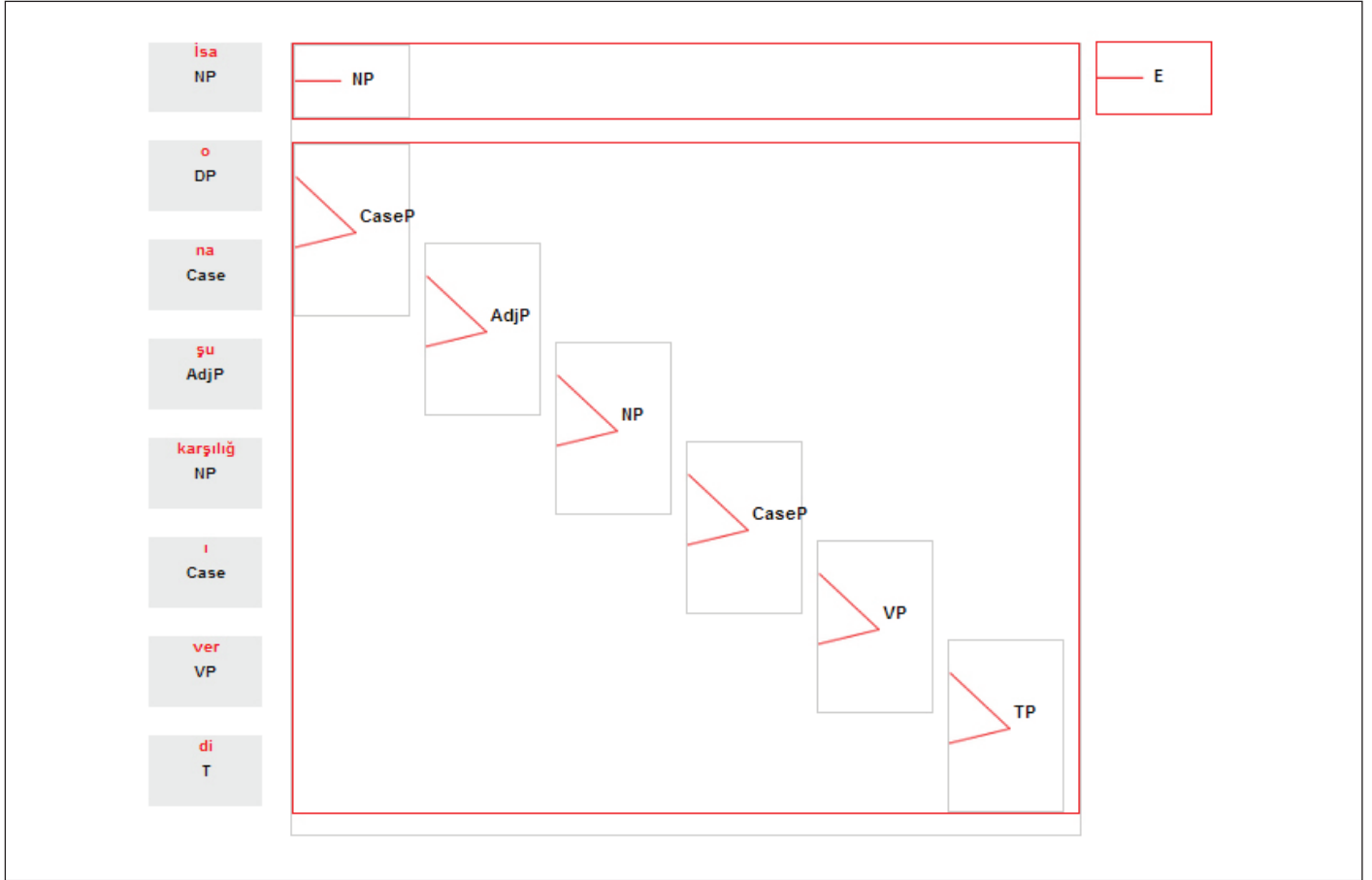
Delete label

Delete label

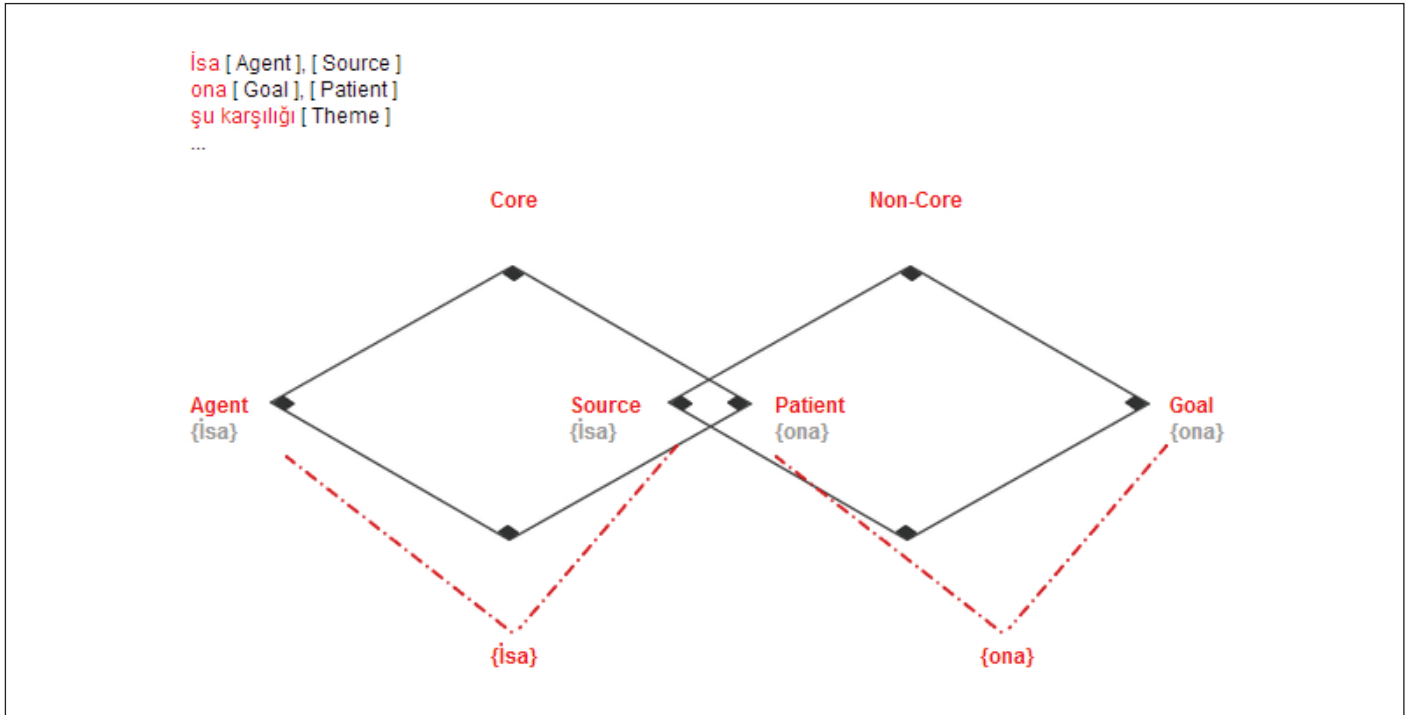
Please Select

**Şekil 2:** Dizimsel/biçimsel-tematik rol işaretleme ara yüzü.

<sup>2</sup>Arayüze <http://www.yelda1367.com/member/login> adresi üzerinden erişilebilir.



Şekil 3: İşaretlenmiş örnek cümlenin otomatik olarak çıkarılmış kural ağacı.



Şekil 4: İşaretlenmiş örnek cümlenin otomatik olarak çıkarılmış latis diyagramı.

tanımını içermektedir ve geleneksel sözlüklerden farklı olarak çok miktarda bilgi sağlamaktadır (Fillmore, Johnson & Petruck, 2003).

Bu veri tabanı, sözcüklerin anlamları altında cümledeki her bir ögenin biçimsel/dizimsel ve anlamsal gösterimlerini tanımlayan karakteristik özellikleri kapsamaktadır.

Şekil 3’de derlemden seçilen örnek bir cümle için kural ağacının otomatik olarak çıkarılmış görüntüsü verilmiştir. Şekil 4’de ise seçilen örnek cümleye ait öğelerin tematik rollerine ulaşmak için otomatik olarak çıkarılmış latis diyagramı gösterilmiştir.

### SONUÇLAR

Bu çalışma ontolojilerin hizalanmasına yönelik bir ön çalışma olarak ele alınırsa, diller arası çeviri ve entegrasyonu sağlama-sına yönelik katkısı açısından önem kazanmaktadır. Türkçe üzerine çalışan araştırmacılara ileri düzeyde veri tabanına erişme imkânı verir. Özellikle Türkçe’nin eklemeli ve özne-nesne-yüklem dizilişine sahip olmasından dolayı makine öğrenmesindeki ve çevirisindeki güçlükleri ortadan kaldırmayı aynı zamanda anlamı belirsiz sözcükleri anlamlı hale getirmeyi sağlar.

Bu çalışma, her bir cümle ögesi için bir işaretlemidir. Bu teknikte gelişmiş ontoloji tabanlı kavram latisleriyle, Türkçe cümlelerin öğeleri tanımlanır. İlerleyen çalışmalarda oluşturulan bu latislerle, incelenen cümle öğelerinin biçimsel/dizimsel özellikleri de dikkate alınarak tematik rolleri üzerinden anlamsal rollerine ulaşılarak sözcüklerin formüllerle hesaplanan olası dağılımlarından sözcük grupları oluşturulabilir.

Bu anlamda, bu çalışma ve beraberinde gelecek çalışmanın aşamaları ile iki dilli ontoloji hizalanması (“bilingual ontology mapping”) sağlanarak, elektronik ortamda geniş kapsamlı Türkçe sözlük yaratılıp İngilizce sözcük birimlerinin Türkçe cümlelerdeki doğru karşılıkları bulunarak makine çevirisi gerçekleştirilmiş olunur.

Bu çalışmaya uygulama açısından bakıldığında, biçimsel ontolojilerin bilgi bankası görevi gören teknik bir araç olmalarından dolayı kullanıldıkları bir diğer çalışma alanının ise eğitim alanı olduğu görülür. Örneğin eğitsel içerik için tasarlanmış bir ontoloji, bir materyalin öğretimsel fonksiyonunu tanımlamada kullanılabilmesi gibi, ek olarak, etkinliğin temel aldığı eğitsel kuramları belli oranda tanımlamada da kullanılabilir. Ayrıca ontolojiler, birtakım çıkarım araçları ve genel bilgi sunum teknikleriyle kullanıldığında kütüphanecilik alanında ve/veya bilgi paylaşımı ile ilgili çalışmalarda da kullanılırlar.

### KAYNAKLAR

- Dahiya, Y.V. (1995). *Panini as a linguist: Ideas and Patterns*. Delhi, India: Eastern Book Linkers.
- Dowty, D. (1991). Thematic Proto-Roles and Argument Selection. *Language*, 67, 547-619.
- Fillmore, C.J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences*, 280, 20-32. DOI: 10.1111/j.1749-6632.1976.tb25467.x
- Fillmore, C. J., Johnson, C.R., & Petruck, M.R.L. (2003). Background to Framenet. *Intertational Journal of Lexicography*, 16 (3), 235-250.
- Ganter, B., & Wille, R. (1996). *Formale Begriffsanalyse: Mathematische Grundlagen*. Berlin: Springer.
- Gildea, D., & Jurafsky, D. (2002). Automatic Labeling of Semantic Roles. *Computational Linguistics*, 28(3), 245-288.
- Jackendoff, R. (1993). On the Role of Conceptual Structure in Argument Selection. A Reply to Emonds. *Natural Language and Linguistic Theory*, 11, 279-312.
- Kılıçaslan, Y., & Güner, E.S. (2011). *Filtering Machine Translation Results with Automatically Constructed Concept Lattices*. Proceedings of 8<sup>th</sup> International Conference on Concept Lattices and Their Applications (CLA 2011), Nancy, France.
- Kasper, S. (2008). *A comparison of ‘thematic role’ theories*. Unpublished master thesis. The Philips University of Malburg.