

KAVRAMSAL DİZİN (THESAURUS) : YAPISI, KULLANIMI VE HAZIRLANMASI

Aytaç YILDIZELİ*

ABSTRACT

Thesauri of scientific terminology are indispensable tools for purposes of information retrieval services and indexing. The extent of a thesaurus in an area of scientific knowledge should depend upon a number of decisions concerning the type; amount and level of information on one side, and user qualifications and financial resources available on the other. Bearing these constraints in mind thesaurus construction can be achieved by employing special methods which lead to special structures.

Thesaurus preparation according to ISO Standard 2788 has been reviewed as an aid to probable designers of Turkish thesauri.

ÖZ

Bilimsel terminolojiye özgü kavramsal dizinler, bilgi erişim hizmetleri ve indekslemede vazgeçilmesi mümkün olmayan araçlardandır. Bir bilim alanında oluşturulacak kavramsal dizinin kapsamı, bir taraftan erişilecek bilginin türü, miktarı ve düzeyine ve öte yandan kullanıcı özellikleriyle parasal kaynaklara bağlıdır. Bu kısıtları gözönünde bulundurarak yapılan bir kavramsal dizin hazırlığı için özel yöntemler kullanmak gerekir ki bunlar da özel dizin yapılarının oluşmasını sağlarlar.

Bu çalışmada, Türkçe kavramsal dizin hazırlayacaklara yardımcı olmak üzere, 2788 ISO Standardlarına dayalı kavramsal dizin hazırlama yöntemi özetlenmiştir.

I. GİRİŞ

Araştırma alanlarındaki uzmanlaşma, araştırma etkinliklerinin süratle örgütlenmesi, belge sayılarının bilimsel ve teknik gelişme doğrultusunda yüksek rakamlara ulaşması ve yayın türlerinin hızla artışı dünyadaki bilimsel belge sayısında büyük bir tırmanmaya neden olmuştur. Dolayısıyla bilgi / belge merkezleri; bilimsel ve teknik konulardaki mevcut bilginin analizi, depolanması ve erişiminde, zamanın en verimli bir biçimde kullanıl-

* TÜRDOK Bilgi Duyuru Birimi Uzmanı

masını amaçlayan yeni yöntemler geliştirmeye yönelmişlerdir. Özellikle, bu merkezlerde kullanıla gelen sınıflandırma sistemlerinin, yeni kavramlar ve konular eklenerek genişleyip büyümesiyle aşırı karmaşıklaşması ve mevcut sınıflandırma sistemlerinin, indekslemede (dizinleme) «karanlık noktalar» denilen daha özel konulara erişimindeki yetersizliği, kavramsal dizinin bir bibliyografik gereç olarak kullanımını zorunlu hale getirmiştir. (1)

Kısacası, bugün terimlerin birbirleriyle olan anlam ilişkilerini gösteren kavramsal dizin, bilginin depolanma ve erişiminin sağlanmasında kullanılan en önemli araçlardan biridir.

Kavramsal dizin yaklaşımı, ilk olarak 15. yüzyılda Feter Mark Roget'in sözcüklerden anlamlara doğru gitmeyen, aksine anlam grupları altında sözcükleri sınıflandıran bir nevi tersine çevrilmiş sözlüğünde görülmektedir.

Roget bu eserine, Yunanca «Hazine» anlamına gelen «Thesaurus» adını vermiştir. (2) Roget'te sadece eşanlamlı sözcüklerin sınıflandırıldığı kavramsal dizin, bilgi erişim ve deplama amaçlarına yönelik olarak kullanılmaya başlayınca, kapsam açısından genişlemiş ve sözcükler arasındaki hiyerarşik ilişkileri belirtme işlevini de yükümlenmiştir.

Bugünkü anlamda kavramsal dizin tanımı, ISO 2788; «Documentation Guidelines for The Establishment and Development of Monolingual Thesauri» da, işlevsel ve yapısal açılarından şu şekilde verilmektedir.

İşlevsel olarak, kavramsal dizin, indeksçiler (dizinciler) ve bilgi kullanıcıları için, belgelerde kullanılan dili daha sıkıştırılmış özel bir sistem diline uyarlamak amacıyla oluşturulmuş terminolojik bir denetim aracıdır.

Yapısal olarak; kavramsal dizin, belirli bir bilim dalında, anlam ve cins yönünden birbirleriyle, ilişkili sözcüklerin denetimli ve dinamik bir terimler listesidir. (3)

«Bilişim Terimleri Sözlüğü» kavramsal dizini «Gömü» başlığı altında «Bir bilgi erişim dizgesinde, derlemin kapsamına giren kavram ve konuları belirten anahtar sözcükleri, bunlar arasındaki anlamsal ve üreysel ilişkilerle birlikte, devirgen bir biçimde içeren bir söz dağarı, ya da sözlük» diye tanımlar. (4)

II) KAVRAMSAL DİZİN HAZIRLAMA

Bir kavramsal dizinin geliştirilebilmesi için harcanacak çabanın niteliği, aşağıdaki kistaslara bağlıdır :

1) Konunun boyutları ve karmaşıklığı :

Kavramsal dizini oluşturma çalışmalarına başlamadan önce, dizine esas alınacak alanlardan hangilerinin daha derinlemesine işleneceğine karar verilmelidir. Böylece, örneğin, «Türkiye Florası Kavramsal Dizini» için kutuplarda yetişen bitkilere ait terimlerin ayrıntılı bir biçimde işlenmesine gerek olmadığı ortaya çıkacaktır.

2) **Yayın türü** : Bilginin depolanması esnasında indekslenecek yayınların türünü bilmek önemlidir. Örneğin, süreli yayınlar için gerekli olan ayrıntı düzeyindeki indekslemeye, kitaplar için gereksinme duyulmayabilir.

3) **Yayın miktarı** : Bilgi / belge merkezlerinde mevcut yayın miktarı ile, ilgilenilen alana ilişkin yayınların ilerideki gelişme hızının tahmin edilebilmesi, hazırlanacak kavramsal dizinin kapsamını saptamada önemli rol oynar.

4) **İndekslemedeki ayrıntı düzeyi** : İndeksleme, kuşkusuz, dayandığı kavramsal dizindeki tanıtılardan daha ayrıntılı yapılamaz. Bu nedenle, tanıtıcı seçimi, indekslemede istenilen ayrıntı düzeyine uyarlanmalı ve böylece hem indeksleme güvenilirliğinin yüksek tutulması ve hem de bilgi-ve emek yönünden önemli tasarruflar sağlanabilmelidir.

5) **Bilgi dizgesinin tipi** : Bilgi dizgesinin, belli bir zaman kesidi içinde geriye dönük bilgi erişim hizmeti mi verdiği, yoksa kullanıcılarına SDI (Selective Dissemination of Information) yani seçimli duyuru hizmeti mi sağladığı, kavramsal dizin tasarımı açısından önceden bilinmelidir. Çünkü SDI hizmeti veren bilgi / belge merkezlerinde oluşturulan kavramsal dizinler, daha önce oluşturulan kavramsal dizinlere bağımlı olmaksızın geliştirilebilir.

6) **Dizge kullanıcıları** : Bilgi dizgesinin, kalifiye bir bilgi uzmanınca mı, yoksa daha geniş bir kullanıcı kitlesi tarafından mı kullanılacağına bilinmesi gerekmektedir.

Çünkü, örneğin kalifiye bir eleman olmaksızın, genel olarak kullanıcılar tarafından çalıştırılacak bir dizge için geliştirilecek bir kavramsal dizin, basit dilde ifade edilmiş, az zaman alıcı ve mümkün olduğunca karmaşıklığı önleyen tanıtıcıları içermelidir.

7) **Kullanıcı sayısı** : Dizgeyi kullananların niteliğinin yanı sıra, niceliğinin de saptanması lazımdır. Ayrıca dizgeyi düzenli olarak gereksinim duyanların miktarının bilinmesi, dizgenin geliştirilme olanaklarının önceden belirlenmesi yönünden önem taşır.

8) **Soru tipleri (Kullanıcıların dizgeye sordukları soruların tipleri) :**

Dizge kullanıcılarının sordukları soruların, çoğunlukla genel anlam-

da sorular mı, yoksa kesin cevabı gerektiren sorular mı olacağı saptanmalıdır. Bu hususun bilinmesi de, indeks dilinin oluşturulmasına esas alınacak kavramsal dizinin kapsamını belirlemek bakımından önemlidir.

9) **İndeksleme çeşidi** : Kavramsal dizini oluşturmadan önce, sözcük gruplandırılması için ne çeşit bir düzenleme yapılacağına karar verilmesi gerekir. Terimleri Boole mantığı ile, araştırma aşamasında ilişkilendiren bir «post-coordinate» indeksleme yöntemi seçilebileceği gibi, sözcüklerin indeksleme esnasında birleştirilmesini engören bir «pre-coordinate» indeksleme yöntemi de seçilebilir. Genellikle post-coordinate düzenleme, bilgisayar teknolojilerinden ve ışıklı kartlardan yararlanıldığında, «Pre-coordinate» düzenleme ise geleneksel kataloglarda veya basılı indeksler için kullanılmalıdır.

10) **Parasal kaynaklar** : Kavramsal dizin hazırlama, çok pahalı projelerden olduğu için, büyük parasal kaynaklar gerektirebilir. Bunun yanı sıra, çalışacak personel ve eldeki donanımın niteliği ve ilgili tüm kuruluşların eğüdüümü, önemli önkoşullardandır. (5)

III) KAVRAMSAL DİZİNİN BDE DİZGESİNDEKİ ROLÜ

Kavramsal dizin geliştirilmeden önce, BDE (Bilgi depolama ve erişim) dizgesinin üst örgütü veya iletişim ağı içindeki rolü ve işlevleri belirlenmiş olmalıdır. Çünkü, kavramsal dizinin amacı, BDE dizgisini kullananların özel gereksinimlerine, görüşlerine, tercihlerine yanıt verebilmesidir. Ayrıca yapısı itibariyle, dizin BDE dizgesinin kullanım sıklığına, dermenin büyüklüğüne ve geri çağırma tekniğine de uygun olmalıdır. Öte yandan, bir BDE dizgesinin başarısı, çok kısa sürede, doğru ve ekonomik bir şekilde bilgiye erişimi sağlayabilmesiyle ölçülebilir. Hiyerarşik ve diğer ilişkileri iyi belirlenmiş, konu kapsamı iyi saptanmış, yanlıklıklar ve belirsizliklerden uzak tanıtıcıları içeren ve iyi bir dil ve düzen içinde sunulan kavramsal dizin, indeksleme, eldeki belgeleri kolayca analiz etme olanağı sağlayacak ve araştırmacının istediği konudaki bilgiye en kısa zamanda erişmesine yardımcı olacaktır.

IV) KAVRAMSAL DİZİNİN YAPISI

Terimlerin kavramsal dizin içinde yerleş ve birbirleriyle ilişkilendirme biçimleri, kavramsal dizinin yapısını oluşturur. Kavramsal dizinde karşılıklı göndermeler ve hiyerarşik ilişkilerle adeta bir ağ örülür.

Bir kavramsal dizin, belgelerin indekslenmesi ve geri çağırılmasında kullanılan bir dizi tanıttıcı (descriptor) içerir. Bu «tanıtıcı dizisine» «İndeksleme dili = indexing language» denir. Bir tanıtıcı, kavramsal dizinde, in-

deksleme işlemi için kullanımına izin verilen bir terim veya birden fazla sözcükten oluşan bir ibare (phrase)dir. Aynı anlayış çerçevesinde tanıtıcı bir kısaltma veya bir simge de olabilir. Bîçimi ne olursa olsun, tanıtıcın işlevi, belgelerdeki kavramları açık bir şekilde tanımlamaktır.

Tanıtıcıların birbirleriyle ilişkileri üç şekilde olmaktadır. Bunlar;

a) Eşanlamlılık (eşitlik) ilişkisi; b) Hiyerarşik ilişki c) İlgili ilişkisidir.

Bir kavramsal dizinde, kullanıcısının ilk aklına gelen sözcük, tanıtıcı olsun veya olmasın en azından şu göndermeler bulunmalıdır.

BT (Broader term)	: Geniş terim kapsamlı tanıtıcı Sözcük konusu tanıtıcıdan daha genel
NT (Narrower term)	: Dar terim Söz konusu tanıtıcıdan daha dar kapsamlı tanıtıcı
RT (Related term)	: İlgili terim Eşanlamlı olmamakla birlikte sözcük konusu tanıtıcıyla genellikle aynı bağlamda kullanılan terim
USE	: Tercih edilen terim
UF (Used for)	: Yerine kullanılan terim (USE gönderiminin zıddı)

Bir kavramsal dizinde, ilişkiler belirtilirken, İngilizce'den türetilmiş USE, UF, BT, NT, RT sembollerinin kullanılması uluslararası standartlaşmanın gereğidir.

Belirtilen ilişkileri kısaca açıklayalım.

a) Eşitlik (Eşanlamlılık) ilişkisi : USE (Kullan) USED FOR (Yerine kullanılmış)

USE Simgesi :

- 1) Bir terimin, tercih edilen anlamdaşını belirtmek için,
- 2) Aynı yazılan ve telaffuz edilen, fakat birden fazla anlamı olan sözcükler (homonyms) için,
- 3) Anlamı aynı olan ve aynı şekilde telaffuz edilen fakat değişik yazılan sözcükler (homophones) için,
- 4) Aynı yazılan ve eşanlamlı olan, fakat değişik telaffuz edilen sözcükler (homographs) için,
- 5) Belirli bir kavramı belirtirken, özel bir terimden daha genel bir tanıtıcı göndermek için,

6) Bir kavramı, iki veya daha fazla sözcükle, anlamsal açıdan bölümlere ayırmak için.

7) İndeksleme (dizinleme) ve erişimde kolaylık sağlamak amacıyla, yarı anlamdaş sözcükleri ilişkilendirmek için.

8) Devrik anlatımları, doğal dil düzenine göndermek için.

9) Gündelik konuşmada kullanılan sözcükleri, bilimsel sözcüklere göndermek için.

10) Zıt terimleri biraraya toplamak için.

11) Kısaltmalardan, terimlerin açık isimlerine veya terimlerden onların kısaltmalarına göndermek için.

12) Kavramsal bir bütünün değişik parça ve derecelerini bir araya toplamak için kullanılır.

USE'ü kullanırken aynı kavramı ifade eden eşanlamlı birkaç terim olduğu zaman, bunlardan sadece bir tanesinin tanıtıcı olarak seçilmesi gereği unutulmamalıdır. Ayrıca eşanlamlılık ilişkisi kurulurken, anlamdaş ve yarıanlamdaş terimler arasındaki farklılık daima gözönünde bulundurulmalıdır. Bu açıdan anlamdaş terimler, hemen hemen aynı anlamı içeren sözcüklerdir. Yarıanlamdaş terimler ise bilim dilindeki anlamları birbirlerinden farklı oldukları halde, kullanılacak bilgi dizgesinin amaçları doğrultusunda eşanlamlı kabul edilen sözcüklerdir. Böylece USE bir bakıma, belgelere daha kolay erişimi sağlamak açısından, bir kavramı anlatmak için kullanılan tüm terimlerin bir araya toplanması amacını da gütmektedir.

A USE B veya B USED FOR A, kullanıcıyı değişik yollardan aynı noktaya götürür. Bunlardan USE sembolü seçeneklerden «yeğlenen» «tercih edilen» terimi belirler. Bu durumda A USED FOR B işaretiyle de yeğlenen terimden diğer terimlere gönderme yapılacaktır.

Şayet, bilgi bankasında kullanılmak üzere tanıtıcılara, örneğin, «Bilgi dizgeleri» yerine C236917 kodu gibi numaralar verilmişse, bu durumda «USE»a gerek yoktur; çünkü böyle sistemlerde A ve B aynı numaraları taşıyacaktır.

b) Hiyerarşik ilişkiler : BT, NT (Geniş ve dar terimler)

Genel/özel ilişkisini tanımlamak ve bunları sergileyebilmek, kavramsal dizin oluşturmanın en zor yanıdır. Çünkü genel ve özel arasındaki hiyerarşi, her zaman tam tartışma gereksindirmeyecek derecede açık olmayabilir. Ayrıca bakış açısına göre de değişebilir.

Kavramsal dizinler oluşturulurken, hiyerarşinin seçilebilir, olması en azından niteliğinin şüpheli olduğu durumlarda ise nesnel tercihler yapılması gerekmektedir. Böyle tercihler yapılırken kullanıcı alışkanlıklarının göz önüne alınması pratik çözümler sağlayabilir. Diğer taraftan araların-

daki hiyerarşik ilişkinin kuşku uyandırdığı terimleri çoğu kez USE bağlan-
tısıyla ilişkilendirmekte yarar vardır.

Kavramların üst ve alt kademelerdeki ilişkilerini belirleyen hiyerarşik
düzen, özel ile genel arasında olduğu gibi parça ile Bütün arasında da
gerçekleşebilir.

Bir kavramsal dizinde hiyerarşik ilişkiler, BT (Broader Term) geniş te-
rim ve NT (Narrower Term) dar terim karşıt sembolleriyle açıklanır. Bu-
rada BT üst kademede tanıtaçın ve NT onun altındaki tanıtaçın görece
yerlerini belirler.

c) İlgili ilişki: : RT (ilişkili terimler)

- 1) Bir terimin anlamca karşıtını belirtmek için,
- 2) Kavramların, üst kademe kavramla bir dallanma ilişkisi içerisinde
bulunması halinde, sözkonusu kavramları bağlantılamak için,
- 3) Akrabalık ilişkisini belirlemek için,
- 4) «Eğitim» ve «Öğretim» arasında olduğu gibi, çoğu kez birlikte kul-
lanılan kavramları ilişkilendirmek için,
- 5) Neden-sonuç ilişkisini yansıtmak için,
- 6) «Ulaştırma» ve «taşitlar» arasında olduğu gibi malzeme, araç vb.
ile bunların işlevleri arasındaki ilişkiyi yansıtmak için,
- 7) Bir türün çeşitleri arasındaki benzerliği göstermek için kullanil-
mektedir.

Eunlardan başka, kavramsal dizinlerde sık kullanılan SN sembolü
dikkati çekmektedir. SN (Scope Note) «kapsam notu», tanıtaçların altında
yer alır ve ait olduğu terimin indeksleme dilindeki anlamını açıklayan ve
sınırlarını çizen tanımları içerir.

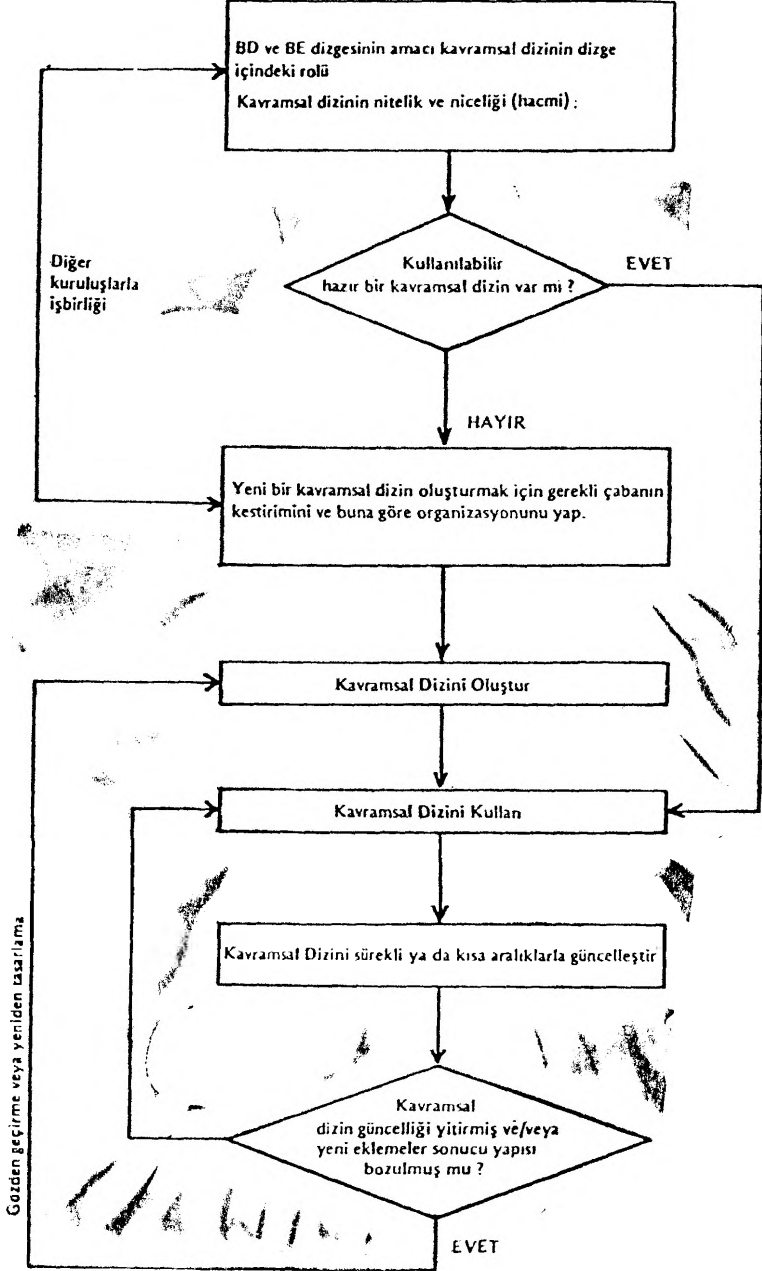
V) KAVRAMSAL DİZİN HAZIRLAMA SÜRECİ

Yeni bir kavramsal dizin oluşturulurken, bilgi deplama ve erişim diz-
gesinde, aynı konuda kullanılabilecek başka bir kavramsal dizin olup ol-
madığının araştırılması gerekmektedir. Şayet hazırlanacak konuda, bir
kavramsal dizin mevcut olup da, üzerinde küçük değişiklikler yapılmak
suretiyle bilgi dizgesinin yeni gereksinimlerine cevap verebiliyorsa, onu
kullanmakta yarar vardır.

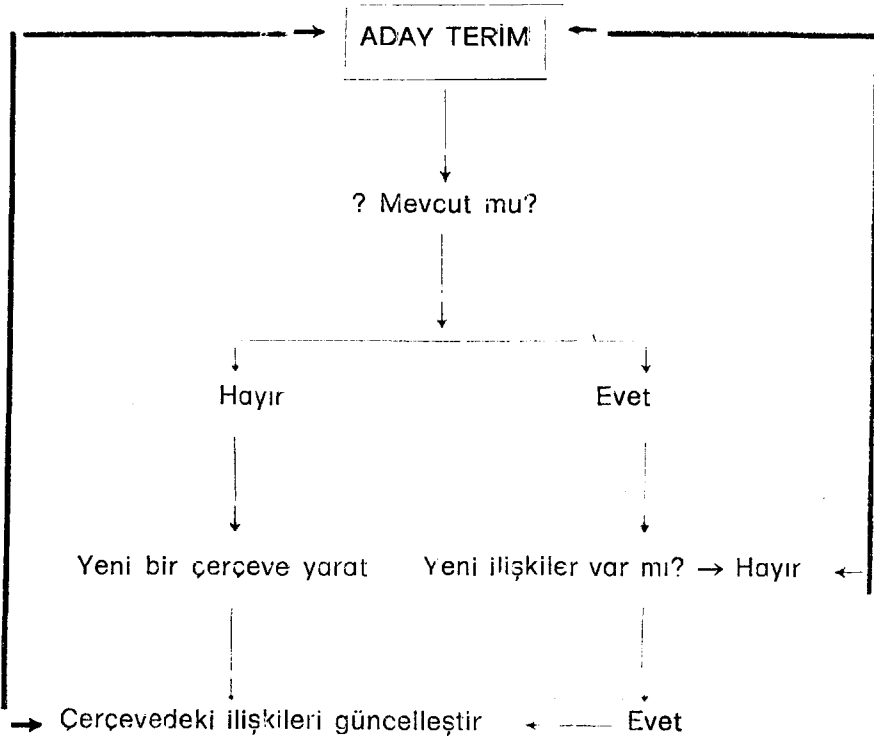
Kavramsal dizin oluşturmada, üç güçlük ortaya çıkabilir. Birincisi, bu
işle uğraşan kişinin kendisini dilci olarak görmemesidir. Gerçi, kavramsal
dizinin malzemesi kavram ve sözcüklerdir, yani dildir. Ancak kavramsal
dizini geliştiren kişinin dilci olmasına gerek yoktur. Sadece kendi konu-
sundaki kavramları ve kullandığı dilin yapısını bilmesi yeterlidir.

İkinci güçlük, seçilen tanıtaçlarla, belge içeriklerinin, kullanıcının ter-

Şekil 1
Kavramsal Dizin Oluşturmada Dikkate Alınacak Noktalar (Akış Şeması)



cihleri ile uyuşum sağlayıp, sağlamayacağıdır. Bu sorunun boyutları, özellikle Türkçe hazırlanacak kavramsal dizinler açısından oldukça geniştir. Çünkü, «yeğlenen terimlerin» birçok dilde bir sözlük taraması ile halledilebilecek bir husus olmasına karşın, Türkçe'de durum daha değişiktir. Örneğin belli bir kavrama ilişkin olarak bilim adamlarının yazılarında kullandıkları terimlerle, Türk Dil Kurumu'nunki gibi yetkili sözlüklerde rastlananlar her zaman aynı değildirler. Türkçe'de eski ve yeni dil gibi sorunların henüz üstesinden gelinmemiş olması, kavramsal dizin oluşturmayı güçleştirmektedir. Böyle güçlüklerin yenilmesi, bir taraftan yılların verdiği indeksleme deneyimine ve kullanıcıların bilgi erişim başvurularının niteliğine ve öte yandan kavramsal dizin bittikten sonra izlenecek kullanım durumunun gerekli kıldığı düzeltmelerin yapılmasına bağlıdır.



Şekil II

Terim İlişkileri Çalışması Akış Şeması

Uçuncü sorun, disiplinlerarası konularda, hazırlanacak kavramsal dizinlerde hazırlayan ekibin, içerilen her anda yeterli konu bilgisine sa-

hip olmayışıdır. Bu sorun, genelde, konu üzerinde uzman kimselerin danışman olarak kullanılmasıyla halledilebilmektedir.

Şekil I, bir kavramsal dizinin oluşturulması ve sürekli yenilenişinin sağlanmasındaki aşamaları bir akış şeması halinde özetlemektedir. Şekil II ise bir kavramsal dizindeki tanıtaçların ilişkilendirilmesi sürecini yansıtmaktadır.

Bunun dışında, tanıtaçlar için yapılacak sözcük seçiminde de uyulması gerekli bazı ilkeler vardır. ISO 2788 standardı bunları şu şekilde belirlemiştir.

a) Mümkün olan her durumda isim kullanılması gereklidir. Kabilse, diğer sözbölükleri, örneğin sıfatlar, isme dönüştürülmelidir.

b) Fiil veya yardımcı fiil kullanılmaz.

c) Sayılabilen terimler için çoğul kullanılması yerinde olur. Yani «kaç tane» sorusuna anlamlı yanıt verilebiliyorsa, terim çoğul olarak alınmalıdır.

d) Tekil terimler ise, genellikle çoğulu bulunmayan veya «ne kadar» sorusuna yanıt veren tanıtaçlar için kullanılmalıdır. Ancak, tekil veya çoğul kullanımı kararı, bir çok durumlarda, yukardaki ilkelerin çerçevesini aşan akıl yürütmeler gerektirebilir.

Genel bir bakış açısından ele alınırsa, Salton bir kavramsal dizinin oluşturulmasına ilişkin olarak şu beş ilkeye uyulmasını önermektedir :

1) Seyrek kullanılan kavramlar alınmamalıdır. Bu durumda belge erişimi ve belge uyuşumunu sağlayacak talep çok az olacaktır.

2) Kullanımı çok yaygınlaşmış terimler alınmamalıdır. Bu durumda bilgi erişiminde konuyla ilgili uygunluk azalacaktır. Buna mukabil, bazı yaygın terimlerden oluşan birleşik terimlerin kullanılması yararlı olabilir.

3) Onemsiz sayılan terimler, elenmiş terimler listesine konulmadan önce dikkatle incelenmelidir.

4) Anlamı açık olmayan terimler mümkünse kapsam dışı bırakılmaktadır.

5) Her kavram sınıfı, yaklaşık olarak eşit kullanım sıklığındaki terimlerden oluşmalıdır. Böylece bir kategorideki terimlerin uyuşum nitelikleri, aşağı yukarı aynı olacaktır. (6)

K A Y N A K Ç A

1. Townley, H. M.; Gee, R. P. **Thesaurus making**. London, Andre Deutch, 1980. 17, 21-22, s.
2. Gilchrist, A. **The thesaurus in retrieval**, London, ASLIB, 1971, 4, s.
3. **ISO 2788-1974. Guidelines for the establishment and development of monolingual thesauri**, Geneva, 1974, 13, s.
4. Köksal, A. **Bileşim terimleri sözlüğü**, Ankara, Türk Dil Kurumu, 1981, 46. s.
5. Aitchison, J.; Gilchrist, A. **A thesaurus construction. A practical manual**. London, ASLIB, 1972, 2-7, s.
6. Soergel, D. **Indexing languages and thesauri : Construction and maintenance**. California, Melville Publishing, 1974, 632. (A Wiley-Becker and Hayes Series Book, Information Sciences Series)