

TÜRKİYE
KÜLTÜR VE SANAT
YILLIĞI

2006

ANKARA ◆ 2006

Dizgi : Osman Bağış

Sayfa Düzeni : Hamza GÜRER

Baskı : Baran Ofset



Elektronik Yayıncılık

Doç. Dr. Mehmet Kara

TÜRK LEHÇELERİ ARASINDA GELİŞTİRİLECEK ÇEVİRİ YAZILIMLARININ GÜÇLÜKLERİ ve KOLAYLIKLARI

Doç. Dr. Mehmet Kara
Fatih Üniversitesi



Yapay zeka ve doğal dil işlemeyle ilgili çalışmalar geliştikçe, çeviri yazılımlarıyla ilgili çalışmalar da gelişme göstermektedir.

Karşılıklı çeviri yapılacak dillerin iç örüntülerinin ve kültür dairelerinin ortaklık düzeyi ne kadar yüksekse, çeviri yazılımlarının başarı oranı da o kadar yüksek olacaktır.

Bu bakımdan birbirlerine uzak dillerle ilgili çeviri yazılımlarının başarısı düşük olmaktadır. Türkçe – İngilizce ve İngilizce - Türkçe çeviri programlarından istenilen verimin alınamamasının asıl sebebi budur. Türkçeye uzak olan bütün dillerde aynı zorluklar yaşanacaktır.

A. Türk Lehçeleri Arasında Geliştirilecek Çeviri Yazılımlarının Kolaylıkları

Aynı kökten gelen ve cümle kurgusu olarak “Özne + Tümleç + Yüklem” şeklinde paralel dizilişe sahip olan aynı kültür dairesine mensup Türk lehçeleriyle ilgili çeviri programları, çok daha kolay geliştirilebilecek ve bunun başarı oranı daha yüksek olacaktır. Şiir dilinde dizilişteki “Özne + Tümleç + Yüklem” paralelliğinin bozulması, yazılım geliştiricileri biraz daha fazla uğraştırabilir.

Türk lehçeleri arasındaki ortaklık sadece cümle dizilişiyle sınırlı değildir. Cümle dizilişi yanında kelime grubu dizilişi, ek-kök sisteminin diziliş sırası, kavram ve kelime hazinesi ortaklığı gibi dil öğeleri bakımından ileri düzeyde ortaklıklar bulunmaktadır.

Özellikle Türkçenin eklemeli bir dil olması ve bütün lehçelerin bu sistemi taşıması, işi kolaylaştıracak önemli bir etken olarak görünmektedir. Bu yapıdan faydalanılarak sadece daha küçük parçalara ayrılamayan en küçük parça olan kök hâlindeki kelimeler, veritabanına girilebilir; kelime gövdeleri ve çekim ekleri biçimbirim çözümlenmeleriyle halledilebilir.

Biçimbirim çözümlenmeleri geliştirilirken fonetik değişkenlik konusu ayrı bir önem taşımaktadır. Çünkü kelime bakımından ortak öğeler çok fazla olsa bile bu ortaklığı örten en önemli etken fonetik kapanmadır. Fonetik kapanmayla kastettiğimiz şudur: Aslında aynı olan kelimenin kök ve ek sisteminde birkaç değişiklik gerçekleşince aynı kelime tanınmaz hâle gelmekte ve zihinlerde ortaklığını yitirmektedir. Fonetik değişkenliğin eklerde köklerden daha fazla olduğunu düşünüyoruz. Çünkü kelime köklerine, kalın/ince/düz/yuvarlak ünlüyle ya da sedalı/sedasız ünsüzle bitmesine göre eklerin bazı lehçelerde en fazla on iki türü gelebilmektedir. Örnek olarak; Türkiye Türkçesinde bir ekin en fazla sekiz türü varken, Kırgızcada on iki türü bulunmaktadır.

Fonetik değişkenlik konusuna şu örnekleri veriyoruz: Türkiye Türkçesindeki *taşlık* kelimesinin Kazakça karşılığı *tastık*, *dinleyip* kelimesinin karşılığı *tımdap*'tır. İlk örnekte olduğu gibi Türkiye Türkçesinde ek başındaki "l", Kazakçada sedasız ünsüzlerden sonra "t"ye dönmektedir: Türkiye Türkçesi "-lık", Kazakça "-tık". Türkiye Türkçesindeki *dinleyip* kelimesi ünlüyle bittiği için "-ip" zarf-fiil ekinde önce araya bir yardımcı ses girmekte, bu kelimenin Kazakça karşılığı da ünlüyle bittiği hâlde zarf-fiil eki doğrudan "-p" şeklinde getirilmekte ve araya bir yardımcı ünsüz girmemektedir: *tımdap*.

Fonetik değişkenlik, sadece kelime köklerine gelen eklerde görülmemekte; ek alan bazı köklerde de buna rastlanmaktadır: Türkiye Türkçesinde *de-* fiili, *dedi*, *demiş*, *desin* örneklerinde *de-* şeklinde iken *diyor*, *diyecek*, *diye* örneklerinde *di-* biçimindedir. Son üç örnekte "y" ünsüzü geniş ünlüyü daraltmıştır. Aynı şekilde *susa-* kelime gövdesi şimdiki zaman eki aldığı anda *susu-* hâlini alıyor. Ancak bu kelime gövdesi şimdiki zaman ekindeki "y"nin tesiriyle ünlü değişmesine uğrarken *susayacak* örneğinde gövdeden sonra "y" geldiği hâlde ünlü değişmesi/daralması gerçekleşmiyor. Bu tür örneklerde sadece kelime kök veya gövdesini veritabanındaki sözlüğe atmak, çevirinin sıhhati açısından yeterli olmayacaktır.

Verdiğimiz örneklerden de anlaşılacağı üzere, lehçeler arasındaki paralellikler ve farklılıkların öncelikle dilciler tarafından bilgisayar diline yatkın bir şekilde işlenmesi gerekmektedir. İşlenmiş malzemeyi esas alarak bilgisayar uzmanları başarımlı yazılımlar geliştirebilirler.

Yukarıdaki örneklerde de görüldüğü gibi, kaynak ve hedef lehçe bünyesindeki kök ve eklerdeki fonetik değişkenlikler konusu işlenirken, herhangi bir kök veya ekin hangi fonetik sistematiği taşıdığı, birleşmeler sırasında bünyesinde değişiklik olup olmadığı; diğer lehçede hangi kök ve ekle karşılandığı iyi tespit edilmelidir.

Konuya ortak birimcikler açısından yaklaştığımızda, bunların düzeyi; Oğuzca, Kıpçakça gibi aynı üst gruba dayanan lehçeler arasında çok daha fazladır. Bu yüzden bir Oğuz grubu Türk lehçesinden diğerine çeviri yapacak bir yazılım, çok daha başarılı sonuçlar verirken; Oğuz grubu Türk lehçelerinin herhangi birinden bir Kıpçak grubu Türk lehçesine çeviri yapacak yazılımın başarı düzeyi daha düşük olacaktır. Daha açık bir ifadeyle, Türkmence ve Azericeden Türkiye Türkçesine yapılacak çeviri yazılımında elde edilen başarı, Kazakça ve Kırgızcadan Türkiye Türkçesine yapılacak çeviri yazılımında sağlanamayacaktır. Öte yandan Kırgızcadan Kazakçaya veya tersine çeviri yapacak bir yazılım da elbette daha başarılı olacaktır.

Yeni alıntı kelimeler açısından ortaklığın düzeyine de bir bakmakta fayda vardır. Arapça ve Farsça kelimeler Türk lehçelerinde çok daha eski ve ortak alıntılardır. Son iki yüzyıl içerisinde Sovyet sistemine tabi olan Türk cumhuriyet ve topluluklarının dilinde Rusça kelimeler yaygınlaşırken, Türkiye Türkçesinde Batı dillerinden alınma kelimeler -özellikle Fransızca ve İngilizce alıntılar- yaygınlaşmıştır. Bu anlamda Türkiye Türkçesi dışındaki lehçelerin birbiriyle ortaklık düzeyi daha fazladır.

Türkiye Türkçesi bünyesinde bulunan yapma kelimelerin bir kısmı, dilin diziliş sistemine uymadığı için diğer Türk lehçeleriyle olan kelime ortaklığına epeyce zarar vermiştir. Çeviri yazılımları için geliştirilecek algoritmalarda, bu açıdan bozulan matematiksel düzenek, bizi biraz uğraştırabilir. Örnek olarak; *bağımsız* ve *ödü* kelimelerini verebiliriz. İlk örneğin sonunda bulunan "-sız" eki, Türkiye Türkçesinin belirgin eklerinden biridir. Fakat kelimenin daha gerisinde bulunan "bağım"daki "-m" yaygın bir şekilde fiilden isim yapma eki olarak kullanılır. Ancak kök görünümlü "bağ", fiil olmayıp isimdir. İkinci örneğin sonunda bulunan "-n" fiilden isim yapma eki olarak kullanılır. Ancak kök tarafında bulunan "ödü" kısmının isim veya fiil olduğu belli değildir. Eğer bu kelime *öde-* fiilinden dilin kurallarına uygun olarak yapılmış olsaydı *öden* olması gerekirdi.

Bu tür kelimelerde hiç biçimbirim çözümlemesi yapmayıp hepsini kökler arasında doğrudan veritabanına atmak belki de en akılcı davranış olacaktır. Bunu yapmak için de kök-ek sistemine göre çözümlenemeyen bütün yapma kelimelerin önceden tespit edilmesi gerekmektedir.

B. Türk Lehçeleri Arasında Geliştirilecek Çeviri Yazılımlarının Güçlükleri

Buraya kadar daha ziyade Türk lehçeleri arasında geliştirilecek çeviri yazılımlarının kolaylıklarından söz ettik. Bu kolaylıklar yanında yazılım geliştirmeyi geciktirecek ve güçleştirecek etkenler de yok değildir.

En büyük eksiklik, taban çalışmalarının yapılmamış olmasıdır. Şu ana kadar çağdaş Türk lehçeleri, çeviri yazılımına uygun bir şekilde işlenmemiştir. Bu konuda ilgili uzmanları eleştirmek bir yere kadar haklı olabilir. Ancak bazı çalışmalar ihtiyaçtan doğar. Bu güne kadar genelde dil ve kültür bakımından birbirimizi anlamaya ihtiyacımız olduğu için buna yönelik çalışmalar yapılmıştır. Bunların da yeterli olduğu söylenemez.

Türk lehçelerini dijital ortamda işlemek için ivedilikle yapılması gereken çalışmaları ortaya koymadan önce şunu belirtmekte yarar görüyoruz: Bu çalışmalar yapılmadan önce bilgisayar uzmanlarıyla Türk lehçeleri uzmanı olan dilbilimcilerin bir araya gelerek nelerin yapılması gerektiğini tespit etmeleri gerekir.

Biz, karşılaşılabilecek güçlükleri ve öncelikle yapılması gerekenleri şöyle sıralayabiliriz:

1. Sözlüklerin Hazırlanması:

Tek tür genel bir sözlük hazırlanabileceği gibi özel sözlükler de hazırlanabilir. Daha verimli sonuç alınabilmesi için genel sözlüklerin yanında özel tipte sözlüklerin hazırlanmasının daha uygun olacağını düşünüyoruz.

Ancak ne tür sözlük hazırlanırsa hazırlansın, şu hususa çok dikkat edilmelidir: Sözlüklerde kelimelerin anlamları, uzun karşılıklar biçiminde değil de mümkünse tek kelime hâlinde verilmelidir. Çünkü çeviri, bir eşleştirme işidir. Ancak tek kelime olarak verilen karşılıklar veritabanına atılarak kolayca eşleştirilebilir.

Bütün sözlükler hazırlanırken karşılıklı iki lehçenin imkânları ve özel durumları göz önünde bulundurulacağından bu sözlüklerin hazırlanmasında karşılıklı olarak her lehçe için, ana dili o lehçe olan kimselerin görev alması gerekir.

Sözlük olarak şunlar hazırlanabilir:

a. Kelime Kökleri Sözlüğü: Böyle bir sözlükte fiil köklerinin geçişli olup olmadığı, hâl eki alıyorsa hangi hâl eki veya eklerini alarak ismi fiile bağladığı gibi ayrıntılar mutlaka verilmelidir. Türkiye Türkçesi için böyle bir sözlük hazırlanırken, hangi fiilin hangi tür geniş zaman ekini aldığı da belirtilmelidir.

Bu sözlükte kökle kaynaşmış yapım ekli gövde de bulunmalıdır: Örnek olarak; *kuç-* fiili aslında “kucaklamak” anlamına gelirken, günümüz Türkiye Türkçesinde bu fiil kök olarak kullanılmadığı için “-ak” fiilden isim yapma eki almış bir gövde olsa bile *kucak* isim gövdesi kök olarak değerlendirilecek ve kök sözlüğüne alınacaktır. *keltür-* fiilindeki “l” ünsüzünün düşmesiyle ortaya çıkan *getir-* fiili de kökünde bu gün *gel-* fiili açık olmadığı için aynı biçimde gövde olarak kök sözlüğüne alınacaktır.

Rahatça tanınabilen köklere gelerek bunlardan gövde oluşturan diğer yapım eklerini almış fiillerin taban kesiti biçimbirim çözümlemesiyle halledileceği için bunların sadece kök kesiti sözlüğe alınacaktır. Ancak bir yapım ekinin diğer lehçede ek olarak karşılığı bulunamazsa, bu tür kelimeler de gövde olsalar bile kök sözlüğüne alınmalıdırlar. Zaten karşılığı olmayan ekler, genelde işlek değildir.

Çeviri yazılımlarında iki lehçe söz konusu olacağı için, Azerice-Türkiye Türkçesi ve Türkiye Türkçesi-Azerice gibi, karşılıklı ikili sözlüklerin hazırlanması gerekmektedir.

Kelime kökleri sözlüğü hazırlanırken bitki ve hayvan adlarıyla yansımali kelimelerin karşılıklarının eksiksiz verilmesine ayrı bir özen gösterilmelidir. Biz, çağdaş Türk lehçeleriyle ilgili metin aktarıırken (çevirirken) en çok bunlardan sıkıntı çektik.

Bitki adları, Türk lehçeleriyle ilgili sözlüklerde tam olarak hangi bitkiyle ilgili oldukları bilinemediği için genelde “Bir tür bitki” şeklinde anlamlandırılmaktadır. Bunun hangi bitki olduğu ve kelime olarak karşılığının ne olduğu sözlükte yer almazsa, çeviri yazılımcısı bunu kodlayamaz ve iki lehçe için kelimeleri eşleştiremez.

Bitki adları kadar sıkıntılı olmamakla birlikte, hayvan adları da lehçe sözlüklerinde ilgili hayvan adının karşılığının tek kelime ile verilmediği durumlarda büyük sıkıntı doğurmaktadır.

Yansımali kelimeler, lehçeler arasında genel kelimelere oranla ortaklık düzeyleri daha az olan dil öğeleridirler. Bu yüzden yansımali kelimelerin anlamlarının tam olarak verilmesi oldukça zordur. Bunların anlamlarının sözlükte net olarak tek kelime ile verilmesi, çeviri yazılımları için büyük kolaylıklar sağlayacaktır.

Bu tür sözlüklerin hepsi, lehçeler arasında karşılıklı/iki bölümlü olarak hazırlanmalıdır. Bunlar hazırlanırken ilgili alanın uzmanlarıyla dilbilimcilerin birlikte çalışması, hem işi kolaylaştıracak hem de hata payını azaltacaktır.

b. **Deyim Sözlükleri:** Bu sözlükler, Türkiye Türkçesi-Kazakça ve Kazakça-Türkiye Türkçesi vb. şekilde iki bölümlü, karşılaştırmalı olarak hazırlanmalıdır.

Bu tip sözlüklerin baş kısmında; ufak ses değişikliği dışında kavram alanı örtüşen ve aynı kelimelerden kurulu deyimler; küçük ses değişikliği ile aynı yapıda kurulu olan, ancak kavram alanı örtüşmeyen deyimler; farklı kelimelerle kurulduğu hâlde kavram alanı örtüşen deyimler, tamamen farklı kelimelerle kurulan ve kavram alanları hiç örtüşmeyen deyimler ayrıntılı olarak ele alınırsa yazılım için büyük faydalar sağlar.

Örnek olarak; Türkiye Türkçesiyle aynı kelimelerden kurulu olan Türkmence *bir çukura tükür*- deyiminin anlamı “bir çukura tükürmek” değil, “ağız birliği etmek”tir. Aslında herkes bir çukura tükürdüğünde aynı hareketi yapmış oluyor. Ancak Türkiye Türkçesinin ses özelliğine göre *bir çukura tükür*- şeklinde ifade edilse bile, deyim şeklinde tam karşılığı verilmedikçe, bu bir anlam ifade etmez.

Türk lehçelerinin kendi içinde deyim sözlükleri bulunmaktadır. Ancak bir lehçeden diğerine deyim deyimle karşılayacak biçimde hazırlanmış karşılaştırmalı sözlükler ne yazık ki bu güne kadar ihmal edilmiştir. Şimdilik bu konuda yegâne ve en kapsamlı sözlük, bir araştırma görevlimle birlikte hazırladığımız *Türkmen Türkçesi-Türkiye Türkçesi Deyimler Sözlüğü*'dür.

Deyimler söz konusu olduğunda, aynı kelimelerin hem tek başına ve gerçek anlamlarıyla hem de deyim birleşimi içinde değişmiş anlamlarıyla metinde bulunabilmesi, çözülmesi çok zor bir problemdir. “Başbakan *sert çıktı*” ve “Aldığı maya *sert çıktı*” cümlelerinde *sert* ve *çık*- kelimeleri ortaktır. “Başbakan *sert çıktı*” örneğinde deyim bulunurken “Aldığı maya *sert çıktı*” örneğinde böyle bir kullanım söz konusu değildir. “Araba *yoldan çıkmış*” ve “Adam *yoldan çıkmış*” cümleleri de benzer örneklerdir. Belki de bu iki örnekteki deyimlerin insanlarla ilgili olması, benzer genel kelimelerin de nesnelere ilgili olması tanımlama için bir ipucu olabilir.

Bu örneklerde de görüldüğü gibi deyim birleşiklerinin hiç deyim olmayan genel kelimelerle kesişme ihtimali yüksektir. Bunların önceden iyi çözümlenmesi gerekir.

c. **Sıklık Sözlüğü:** Bu güne kadar herhalde biricik sıklık sözlüğümüz İlyas Göz tarafından hazırlanan *Yazılı Türkçenin Kelime Sıklığı Sözlüğü* (Türk Dil Kurumu Yayınları, Ankara, 2003) olsa gerektir. Oysa sıklık sözlükleri, bir dilin değişik seviyelerde sözlüklerinin hazırlanmasında önemli bir kılavuz niteliğindedir. Her lehçeye ait sıklık sözlüğünün ivedilikle hazırlanması gerekir. Çünkü geliştirilecek bir

çeviri yazılımında kullanılacak sözlük için hangi kelime köklerinin öncelikle veritabanına yükleneceği, ancak bu sayede bilinecektir.

d. **Özel Sözlükler:** Özel sözlükler olarak şunlar hazırlanabilir:

Terim Sözlükleri: Değişik bilim dallarına ait terimleri içeren sözlüklerin lehçeler arasında karşılıklı olarak hazırlanması, çeviri yazılımında büyük kolaylıklar sağlayacaktır. Ancak bu tür sözlükleri ilgili alanın uzmanlarıyla dilbilimcilerin birlikte hazırlamasında büyük faydalar vardır. Çünkü bunların birçoğu alıntı kelime olduğu için karşılık bulmakta büyük sıkıntılar çekilmektedir.

Yalancı Eş Değerler Sözlüğü: Yalancı eş değerlik, kaynak ve hedef lehçede kullanılan aynı kelimelerden birinin aslına göre çok farklı bir anlam taşımasından kaynaklanır. Ancak bu tür kelimeleri kullanan kişiler, bunlara ait anlamların iki lehçede de aynı olduğunu zannederler.

Örnek olarak bir Azerinin “inmek”, “konuşmak”, “ön” anlamlarında kullandığı *düş-*, *danış-*, *gabak* kelimeleri; bir Türkiye Türkü tarafından “düşmek”, “danışmak” ve “kabak” olarak algılanabilir.

Yine aynı şekilde bir Türkmenin “rüşvet”, “elem” anlamlarında kullandığı *pa-a* ve *besret* kelimeleri; bir Türkiye Türkü tarafından “para” ve “hasret” olarak algılanabilir. Bunun tam tersi de olabilir.

İstem (Valenz) Sözlüğü: İsimlerin hangi hâl eklerini alarak fiillere bağlandığı konusu Türk lehçelerinde pek işlenmemiştir. Ancak çeviri yazılımları için bu konunun karşılıklı iki lehçe dikkate alınarak mutlaka işlenmesi gerekir. İstem sözlükleri, lehçeler arasında karşılıklı olarak hazırlanmalıdır. Bu sözlükte sadece fiil kök ve gövdelerine yer verilmeli ve bunların hangi hâl eklerini aldıkları belirtilmelidir.

Örnek olarak *dol-* fiiline bağlanan isim Türkiye Türkçesinde vasıta hâli eki “-la/-le” alırken, Türkmencede aynı fiile bağlanan isim uzaklaşma hâli eki “-dan/-den”i almaktadır: *Asmaniň ýüzi derrev tozandan doldı*. “Gök yüzü hemen dumanla doldu”.

Türkmencede “anlamak” anlamına gelen *düşün-* fiiline bağlanan isim yaklaşma hâli eki “-a/-e”yi alırken, Türkiye Türkçesinde “düşünmek” anlamına gelen aynı fiile bağlanan isim yükleme hâli eki “-ı/-i, -u/-ü”yü almaktadır: *Bu gün özleriniň yal-nışandıklarına düşündiler*. (“Bu gün hata ettiklerini anladılar.”). *Onlar eski günlerini düşündüler*.

Sesteş Kelimeler Sözlüğü: Çeviri yazılımı yapılacak iki lehçenin sesteş kelimeler sözlüğü, bu iki lehçe arasında karşılıklı olarak hazırlanmalıdır. Böylece bulanıklığa yol açacak sesteşlerin tamamı ortaya çıkmış olacaktır. Bu yapılırsa ve yazılım esnasında tanımlama yapılırsa, bilgisayarın bunları ayırt etmesi biraz daha kolaylaşacaktır.

Sesteş kelimeler sözlüğünün baş kısmında sözlük dışı sesteşleşmeler konusu da ele alınmalıdır. Sözlük dışı sesteşleşme, bazı kök ve eklerin birleşmesi sonucu aynı sesleri taşıyan diğer kelimelerle sesteş hâle gelmesidir.

Eş Anımlı Kelimeler Sözlüğü: Eş anlamlı kelimeler sözlüğü de anlam bulanıklıklarının giderilmesi ve kavram alanlarının tespitinde çeviri yazılımları için vazgeçilmez bir kaynak niteliğindedir. Bu tür sözlükler de lehçeler arasında karşılıklı olarak hazırlanmalıdır.

Eş anlamlı kelimeler sözlüğünün yanı sıra, zıt anlamlı kelimeler sözlüğü de hazırlanabilir.

Buraya kadar üzerinde durduğumuz sözlük çalışmaları yapılmadan sağlıklı çeviri yazılımlarının geliştirilmesi pek mümkün değildir.

Çeviri yazılımını zorlaştıracak başka hususlar da bulunmaktadır. Bunlardan birisi "çok anlamlılık"tır.

2. Çok Anlamlılık: Bir kelimenin çok fazla anlamının bulunması, çeviri yazılımlarının önemli zorluklarından biri olacaktır. Bu durumda veritabanına atılan kelimelerde her anlam, dilin hangi kesitiyle ilgili metne aitse ve bu açıdan tanımlama yapmak mümkün olursa, iş biraz kolaylaşabilir. Hatta çeviri programı, bir metin girilir girilmez bu metin hangi dil kesidine aitse ona yönelecek şekilde tasarlanabilir. Bu kesitler; edebî metin, tıp metni, tarih metni, dinî metin, hukuk metni, teknik konulu metin vb. olabilir.

Bu çabayla çeviri yazılımı veritabanının daha başarılı ve kullanışlı olması sağlanabilir.

Türk Dil Kurumunun ağ sayfasında bulunan "Güncel Türkçe Sözlük"e göre Türkiye Türkçesinde en çok anlamı olan kelimelerden *tut-* fiilinin elli yedi (57), *gel-* fiilinin otuz yedi (37), *at-* fiilinin otuz beş (35), *ayak* isminin on altı (16), *baş* isminin on üç (13), *yol* isminin de on iki (12) anlamı bulunmaktadır.

Kaynak lehçedeki çok anlamlı bir kelime, hedef lehçede aynı kelimeyle ve paralel anlamla karşılık bulursa, iş biraz kolaydır. Ancak başka kelime veya kelimelerle bu anlamın karşılanması gerekirse, uygun kelime ve anlamların seçimi biraz zor olacaktır.

Bu durumda daha çok anlama sahip oldukları için fiillerin birçoğu, çeviri yazılımını biraz çetinleştirecektir. Çok anlamlı bazı isimler de bunlardan geri kalma-yacaklardır.

Çeviri yazılımının bir başka sıkıntılı yanı, sözlük dışı sesteşleşmelerin çözümünde yaşanacak problemlerdir.

3. Sözlük Dışı Sesteşleşmeler: Köklerin eklerle birleşmesi ve metnin kurulmasından sonra ortaya çıkan sesteş kelimeler ile asıl sesteş kelimelerin çözümlenmesinde epeyce sıkıntılar yaşanabilir. Bunlar bir tür biçimbirim keşişmesidir. Bu tür yapılarla biçimbirim bulanıklıklarının giderilmesi için ayrıntılı çalışmaların yapılması gerekir. Aksi hâlde geliştirilen yazılım, sesteş öğeler için hangi biçimbirim çözümlemesini kullanacağını bilemeyecek ve yanlış anlamlandırma söz konusu olacaktır.

Türkiye Türkçesine ait olan aşağıdaki dört cümlenin yüklemine, bütün sesleri aynı olduğu için, sözlük dışı sesteşleşmeye örnek olarak verebiliriz:

Koyunu *yüzdü*.

Suda *yüzdü*.

Toplamı *yüzdü*.

Güzel bir *yüzdü*.

Bu örneklerde sesteş öğelerin ilk ikisi fiil, diğer ikisi isim köküne sahiptir. İlk iki örnekte *yüz-* fiili, bilinen geçmiş zaman ekini almıştır ve teklik 3. şahsa ait hareketi temsil eder. Ancak her iki *yüz-* fiilinin anlamları da farklıdır. Son iki örnekteki isim kökünden sonra *-i* fiili düşmüştür. Her iki örnek de *i-* fiiline gelen bilinen geçmiş zaman ekine sahiptir ve 3. şahsı temsil ederler. Ancak bu isim kökleri, ayrı anlamları taşımaktadırlar.

Bu dört sesteş kelimenin biçimbirim çözümlemesini bilgisayarın ayrı ayrı ve doğru olarak yapabilmesi, birinci aşamadaki zorluktur. İkinci zorluk ise, bunların uygun bir biçimde tek tek anlamlandırılmasıdır.

Türkiye Türkçesi için şu örnekleri de verebiliriz: *yaz-ın* "siz yazın", *yaz-ı-n* "senin yazın", *yaz-ı-n* "yaz mevsiminde", *yaz-ı-n* "edebiyat"; *yaz-ı-cı* "orduda kâtip", *yaz-ıcı* "bilgisayarda yazılmış metnin çıktısını veren araç"; *çık-tı* "o çıktı", *çık-tı* "üretim sonucu ortaya çıkan ürün; artık; bilgisayarda yazılmış bir metnin kâğıda dökülmüş biçimi; mezuniyet belgesi" vb. Bu örneklerin biçimbirim çözümlenmeleri de zaman zaman farklı olmaktadır. Biçimbirim çözümlenmeleri ister farklı, isterse aynı olsun; bunların her biri farklı kavramları ifade ettiğinden anlamlarının ayrıştırılması biraz zor olacaktır.

Kazakçada *barıs* biçiminde iki kelime bulunmaktadır. Bunlardan biri "pars", diğeri de "gidiş, istikamet, yön" anlamına gelir. "Pars" anlamına gelen kelime (*barıs*) kök hâlindeyken, diğeri "varmak, gitmek" anlamlarına gelen *bar-* fiil köküne "-ıs" fiilden isim yapma ekinin getirilmesiyle yapılmıştır.

Yine aynı lehçede iki *aştık* kelimesi bulunmaktadır. Bunlardan biri "açlık", diğeri "biz açtık" anlamına gelir. Birinin kökü isim, diğerininki fiildir. Ancak fonetik değişimler sonucu bu kelimeler sesteş duruma gelmişlerdir.

Türkiye Türkçesi ve Kazakçadan verdiğimiz örnekler, bütün sesleri birebir örttüğü için bunların çeviri yazılımlarında çok özel tanımlamalar gerektireceği ve yazılımı hazırlayanları epeyce uğraştıracığı kesindir. Bu yüzden bulanıklığa yol açacak sözlük dışı sesteşleşmeler üzerinde ayrıntılı çalışmaların yapılması gerekir. Bunlar yapılmadan ve yazılım esnasında tanımlamaları gerçekleştirilmeden bilgisayarın sözlük dışı sesteş öğeleri birbirinden ayırt etmesi ve bu tür kelimeleri doğru anlamlandırması mümkün değildir.

Aynı sesteşleşmeler deyim birleşikleriyle genel kelimeler arasında da görülebilir. Türkiye Türkçesinden bir örnek verelim:

Yüz verdim, elli kaldı.

Yüz verdim, hepsi bundan kaynaklandı.

Bu örneklerden ikincisinde deyim bulunmaktadır. Deyimin bulunduğu kısımdaki kelimeler birleşik olarak algılanacak şekilde tanımlanmalıdır. Birinci örnekteki *yüz* ve *verdim* kelimelerininse ayrı ayrı algılatılması gerekir.

İğde kokusu aldım ("İğde kokusu" adlı esansı satın aldım" anlamında).

İğde kokusu aldım ("İğde ağacı çiçeğinin kokusunu aldım" anlamında).

Son örneğin ikinci cümlesinde, *koku almak* deyimini kullanılmıştır. Bu iki cümlede de görüldüğü gibi biçimbirim çakışması sadece tek kelimedede görülmemekte, birkaç kelimededen ibaret cümlelerde de karşımıza çıkmaktadır.

Lehçe içi veya lehçeler arası kavram alanlarının tam olarak örtüşüp örtüşmediğinin tespiti, çeviri yazılımlarıyla ilgili bir diğer zorluktur.

4. **Kavram Alanı ve Kelime Ailesi İlişkileri:** Kavram alanı; birbirine yakın kavramları, eş anlamlıları kapsayan alandır. Türkiye Türkçesinde kullanılan *gam-keder-tasa-kaygı-elim-üzüntü-hüzün* gibi kelimeler aynı kavram alanı içinde bulunurlar. Bu tür örneklerin kavram alanları çok defa birbiriyle eş düzeyde olmaz. Ör-

nek olarak; *keder* ile *hüzün*'ün kavram alanları tam olarak örtüşmez. Çünkü *keder*'de daha derin bir üzüntü varken, *hüzün*'de daha masum ve sevimli bir üzüntü vardır.

Bu durum hem isimlerde hem de fiillerde karşımıza çıkabilir. Türkiye Türkçesinde kullanılan *seyret-* ve *bak-* kelimeleri eş anlamlı gibi görünmekle birlikte bunların anlamları epeyce farklıdır. *seyret-*'te uzun süreli bakma söz konusuyken, *bak-*'ta daha kısa süreli bir bakış ifadesi vardır.

Alıntı kelimelerde de bu duruma dikkat etmek gerekir. Türkiye Türkçesinde bulunan Arapça alıntı iki kelime olan *surat* ve *suret* örnekleri bu açıdan ilginçtir. Aynı kaynaktan gelmekle birlikte bunların kavram alanları epeyce değişmiştir.

Türkmencede kullanılan *garrı-gartañ-kempir-goca-yaşulı-çal-aksakgal* kelimeleri de genel olarak yaşlı insanı anlatmakla birlikte aralarında farklılıklar bulunmaktadır. Bu örneklerden *aksakgal* ve *yaşulı*, yaşlı kavramıyla ilgili olmakla birlikte tecrübelerinden faydalanılacak bilge yaşlıları temsil ederler.

Türk lehçelerinde buna benzer örnekler epeyce çoktur. Bu yüzden bir kaynak lehçedeki bu tür kelimelerin birbiriyle örtüşme durumları öncelikle ele alınmalı, daha sonra da hedef lehçede bulunan ve aynı kavram alanıyla ilgili olan kelimelerin hangisinin en fazla örtüştüğü tespit edilmelidir. Aksi hâlde kelimeler tam karşılıklarını bulamazlar ve aktarma (çeviri) sağlıklı olmaz.

Bazen de aynı kavram alanıyla ilgili kelimelerin her birinin Türk lehçelerinde tercih edildiği özel yerler vardır. Türkiye Türkçesindeki *ak-beyaz* ve *kara-siyah* kelimelerini bu açıdan inceleyelim. Görünüşte *ak* ile *beyaz*, *kara* ile *siyah* her zaman eş anlamlı gibidirler. Ancak kullanımda her birinin tercih edildiği özel yerler vardır. “*beyaz* peynir” diyor, ancak “*ak* peynir” denmiyor. “*Ak* akçe *kara* gün içindir” diyor da “*Beyaz* akçe *siyah* gün içindir” denmiyor. Çevirilerde bu tür ifadelerle karşılaşılabilir; bunlar önemli çeviri yanlışlarıdır.

Bir örnekle de olsa, kavram alanı açısından iki lehçeyi karşılaştırmakta fayda vardır. Türkiye Türkçesinde *gitmekte* ve *gidiyor*, şimdiki zamanı ifade eden örnekler olarak bilinir. Türkmencede de *gidip dur* ve *gidyär* örnekleri şimdiki zamanı ifade eder. Bunları kendi arasında eşleştirdiğimizde; *gitmekte*'nin tam karşılığı *gidip dur*, *gidiyor*'unki ise *gidyär* olacaktır. Çünkü *gitmekte* ve *gidip dur*, sürekli şimdiki zamanı; *gidiyor* ve *gidyär* ise, normal şimdiki zamanı ifade ederler.

Kelime ailesi üzerinde yapılacak çalışmalar da çeviri yazılımlarının geliştirilmesinde önemlidir. Kelime ailesi, aynı kökten türeyen bütün kelimeleri içine alır. Örnek olarak; Türkiye Türkçesindeki *bilgi*, *bilgili*, *bilgisiz*, *bilgisizlik*, *bilim*, *bilgin*, *bilgiç*, *bilgiçlik*, *bildir-*, *bildiril-*, *bilgilen-*, *bilgilendir-*, *bildiri*, *bilişim*, *bilinç*, *bilinçli*, *bilinçsiz* sözcükleri aynı kelime ailesine mensupturlar.

Kelime ailesinde yer alan kümelerin kök kesiti, yukarıdaki örnekler kadar açık olmayabilir. *bıç-* fiil kökünden türeyen *bıçak*, *buçuk*, *bucak* kelimeleri de aynı aileden olmakla birlikte ses değişimleri neticesinde bunların kökleri; *bıç-*, *buç-*, *buc-* şeklini almıştır.

Benzer bir şekilde *birleşim* ve *bileşik* örnekleri de aynı aile içinde yer alırlar. Bunlardaki kelime kökleri ise *bir* ve *bi* olarak karşımıza çıkar.

Verilen örneklerden *bucak* ve *buçuk*, gerektiğinde biçimbirim çözümlemesine tabi tutulmayıp doğrudan kök sözlüğüne alınabilir. Biçimbirim çözümlemesiyle halledilemeyecek durumda olan benzer örneklerin hepsinde aynı yola başvurulabilir.

Kelime kökleri ve yapım ekleriyle türetilmiş gövdeler kelime ailesinin ana malzemesi olduğu için, eklemeli dillerle ve tabii ki Türkçeye ilgili çeviri yazılımları geliştirmede kelime aileleri üzerinde önemle durulması gerekir.

Kelime ailesi çalışmaları, Türk lehçelerinde bulunan paralel kelime ailelerine ait kavram alanlarının birbirleriyle örtüşüp örtüşmediğinin tespitine imkan tanıyacak ve bu sayede metnin doğru anlamlandırılmasında önemli hizmette bulunacaktır.

Kavram alanı ve kelime ailesi çalışacak kişilerin dilbilgisinin diğer kesitleri yanında özellikle anlambilim, sesbilgisi ve etimolojiyi çok iyi bilmeleri gerekir.

Çeviri yazılımlarıyla ilgili bir diğer zorluk da uygun gramer yazımıdır.

5. Uygun Gramer Yazımı: Sağlıklı çeviri yazılımı geliştirebilmek için matematiksel düzeyde yazılmış gramerlerin hazırlanmış olması gerekir.

Bilgisayara uygun Türk lehçeleri gramerlerinin aktarma (çeviri) yapılacak her iki lehçe de göz önünde bulundurularak karşılaştırmalı olarak yazılması gerekir. Bu gramerler hazırlanırken her lehçe için, ana dili o lehçe olan kimselerin görev almasına ayrı bir önem verilmelidir. Çünkü kavramların ayrıntıları, hayatın derinliklerinde gizlidir.

Gramerler yazılırken biçimbirim çözümlemesinin sağlıklı olarak yapılabilmesi için ek-kök sistemi ile fonetik değişkenlik konusunun ayrıntılı olarak işlenmesi gerekir. Kelime grubu ve söz dizimi kısmında özellikle iki lehçenin farklı yanlarının iyice ortaya konması gerekir. Anlamla ilgili kısımda iki lehçenin anlam bulanıklıklarına yol açacak kısımları dikkatlice ele alınmalıdır. Kök ve yapım eki birleşmelerinde hiçbir farklılık yoksa bu tür gövdeler, biçimbirim çözümlemesiyle halledilmelidir.

Ancak bütün bu çalışmaların, ayrıntılı olarak bilgisayar uzmanının rahatlıkla anlayabileceği açık bir dille ortaya konulması çok önemlidir. Zaten bilgisayar destekli dilbilimi çalışmalarının zor yanlarından biri de dilbilimci ile yazılımcının birbirini anlamakta çektiği sıkıntı olacaktır. Daha önce gerçekleştirdiğimiz alfabe çevirisi yazılımında bunu bizzat yaşadık.

Gramerdeki İnce Ayrıntılardan Birkaç Kesit: Örnek olarak geniş zaman ekinin Türkiye Türkçesinde hangi fiillere ne şekilde geldiği konusu, önceden iyi tanımlanması gereken bir husustur. Bu lehçede “-ar/-er” veya “-r” gelmesinin tam olarak düzenli bir sistematiği yok gibidir: Dar-yuvarlak ünlülü *dur-* fiiline *dur-u-r*, ancak aynı özelliği taşıyan *tut-* fiiline *tut-ar*; geniş-yuvarlak ünlülü *gör-* fiiline *gör-ü-r*, aynı özelliği taşıyan *böl-* fiiline *böl-er* biçiminde gelmektedir.

Oysa bir diğer Oğuz grubu Türk lehçesi olan Türkmencede geniş zaman eki kalın sıradan olan fiillerde “-ar”, ince sıradan olan fiillerde “-er”dir ve oldukça sistematik bir yapıya sahiptir.

TDK tarafından hazırlanan *Türkçe Sözlük*'te ve *Güncel Türkçe Sözlük*'te hangi fiilin ne tür geniş zaman ekini aldığına belirtilmiş olmasını çok yerinde bir yaklaşım olarak değerlendiriyoruz.

Geniş zaman ekinin olumsuzu Türkiye Türkçesinde “-ma + -z”; Azerice ve Türkmencede ise, “-ma + -z” veya “-ma + -r” biçimindedir. Türkiye Türkçesinde birinci şahıslarda “-z” geniş zaman eki de düşmüş ve değişmiş durumdadır: *gelmezem* > *gelmem*, *gelmezüz* > *gelmeyiz*:

Türkiye Türkçesi:	Türkmençe:	Azerice:
<i>gelmem</i>	<i>gelmerem</i>	<i>gelmerem</i>
<i>gelmezsin</i>	<i>gelmersiň</i>	<i>gelmezsen</i>
<i>gelmez</i>	<i>gelmez</i>	<i>gelmez</i>
<i>gelmeyiz</i>	<i>gelmeris</i>	<i>gelmerik</i>
<i>gelmezsiniz</i>	<i>gelmersiňiz</i>	<i>gelmezsiňiz</i>
<i>gelmezler</i>	<i>gelmezler</i>	<i>gelmezler</i>

Yukarıda da değindiğimiz gibi, “kelime kökü + yapım eki + çekim eki” sırasındaki çekirdek dizilişte bulunan matematiksel düzenek, algoritmaların geliştirilmesinde büyük kolaylıklar sağlayacaktır. Ancak bu kısımlarda her zaman paralel yapılar da karşımıza çıkmayabilir. Örnek olarak değişik zaman çekimlerinin sorusunda birinci ve ikinci şahıslarda soru eki şahıs ekinden önce gelirken, aynı ek, çok-

luk üçüncü şahısta, şahıs ekinden sonra gelmektedir: *gülmüş müyüm*, *gülmüş mü-sün*, *gülmüş müyüz*, *gülmüş mü-sünüz*; ancak *gülmüş müler* değil *gülmüşler mi* doğrudur. *gelir miyim*, *gelir misin*, *gelir miyiz*, *gelir misiniz*; ancak *gelir miler* değil *gelirler mi* doğrudur. *bilecek miyim*, *bilecek misin*, *bilecek miyiz*, *bilecek misiniz*; ancak *bilecek miler* değil *bilecekler mi* doğrudur.

Aynı şekilde öğrenilen geçmiş zamanın hikâyesinde birinci ve ikinci şahıslarda hikâye eki, şahıs ekinden önce gelirken, aynı ek, çokluk üçüncü şahısta, şahıs ekinden sonra gelmektedir: *uyanmıştım*, *uyanmıştın*, *uyanmıştık*, *uyanmıştınız*; ancak *uyanmıştılar* değil *uyanmışlardı* doğrudur. Bu gün *uyanmıştılar* şekli de zaman zaman kullanılmakla birlikte ağız özelliği taşır.

Şu örnekte ise değişen diziliş, ekteki düzlük-yuvarlaklık durumunu da etkilemiştir: *görmüştiim*, *görmüştiin*, *görmüştü*, *görmüştüük*, *görmüştiünüz*, *görmüşlerdi*. Diğer bütün şahıslarda bilinen geçmiş zaman eki kendisinden önce gelen “-müş” ekindeki ünlü yuvarlak olduğu için “-tü” biçimindeyken, dizilişin değişmesi ve araya “-ler” ekinin girmesiyle aynı ekin ünlüsü düzleşmiş ve “-di” biçimini almıştır. Öte yandan öğrenilen geçmiş zaman ekinin sonunda bulunan ünsüz (“ş”), sedasız olduğu için diğer şahıslarda bilinen geçmiş zaman ekinin ünsüzü “t” iken, çokluk üçüncü şahısta diziliş değiştiğinden ve “-ler” çokluk üçüncü şahıs ekinin sonundaki “r” sedalı olduğundan bunun arkasından gelen bilinen geçmiş zaman ekinin ünsüzü de sedalılaşılarak “d” olmuştur.

Lehçeler arasında da eklerin diziliş sırası ve fonetik değişkenlikleri farklı olabilmektedir. Örnek olarak; Türkmencedeki şimdiki zaman ekinin kökle birleşmesi esnasında hiçbir fonetik değişkenlik görülmezken, Türkiye Türkçesinde hem ünlü değişimleri olmakta hem de ince sıradan olan kelimelerde kalınlık-incelik uyumu bozulmaktadır:

Türkmence:	Türkiye Türkçesi:
<i>başlayarın</i>	<i>başlıyorum</i>
<i>başlayarsıñ</i>	<i>başlıyorsun</i>
<i>başlayar</i>	<i>başlıyor</i>
<i>başlayarıs</i>	<i>başlıyoruz</i>
<i>başlayarsıñız</i>	<i>başlıyorsunuz</i>
<i>başlayarlar</i>	<i>başlıyorlar</i>
<i>isleyärin</i>	<i>istiyorum</i>
<i>isleyärsiñ</i>	<i>istiyorsun</i>
<i>isleyär</i>	<i>istiyor</i>

<i>isleyäris</i>	<i>istiyoruz</i>
<i>isleyärsiñiz</i>	<i>istiyorsunuz</i>
<i>isleyärler</i>	<i>istiyorlar</i>

Bu iki lehçe soru eki bakımından karşılaştırıldığında; Türkiye Türkçesinde kelimededen ayrı, Türkmencede ise kelimeye bitişik yazılan soru ekinin geliş düzeninde de farklılıklar bulunmaktadır:

Türkmençe:	Türkiye Türkçesi:
<i>baryarınmı</i>	<i>gidiyor muyum</i>
<i>baryarsıñmı</i>	<i>gidiyor musun</i>
<i>baryarmı</i>	<i>gidiyor mu</i>
<i>baryarısım</i>	<i>gidiyor muyuz</i>
<i>baryarsıñızım</i>	<i>gidiyor musunuz</i>
<i>baryarlarmı</i>	<i>gidiyorlar mı</i>
<i>gelyärinmi</i>	<i>gidiyor muyum</i>
<i>gelyärsiñmi</i>	<i>gidiyor musun</i>
<i>gelyärmi</i>	<i>gidiyor mu</i>
<i>gelyärisim</i>	<i>gidiyor muyuz</i>
<i>gelyärsiñizim</i>	<i>gidiyor musunuz</i>
<i>gelyärlermi</i>	<i>gidiyorlar mı</i>

Şimdi Türkmençe ile Türkiye Türkçesindeki şimdiki zamanın normal çekimlerini ve soru çekimlerini örnekleyerek bu yapının çeviri yazılımına uygun bir tanımlamasını yapmaya çalışalım:

Türkiye Türkçesinde Ek: -yor.

Türkmencede Ek: -yar/-yär.

Türkiye Türkçesindeki Fonetik Değişkenlikler: Türkiye Türkçesinde “-yor” biçimindeki şimdiki zaman eki, tek tip ve kalın sıradan ünlü bulundurduğundan ince sıradan ünlü almış fiil kök veya gövdelerine geldiğinde kalınlık incelik uyumu ortadan kalkar. Bu yüzden fiilin şimdiki zaman ekine kadar olan kısmı ince sıradan olsa bile, zaman ekinin ünlüsü sadece kalın sıradan geldiği için şimdiki zaman ekinden sonraki bütün ekler kalın sıradan getirilmek durumundadır: *biliyor musunuz*.

“y” sesi, inceltici özelliğinin yanı sıra daraltıcı etkiye de sahiptir. Bu yüzden şimdiki zaman ekinin başında bulunan “y” ünsüzü, kendisinden önce gelen geniş ünlüleri daraltarak *başlayıyor* > *başlıyor* ve *istiyor* > *istiyor* örneklerinde görüldüğü gibi “a”dan “ı”ya ve “e”den de “i”ye çevirir.

Türkmencedeki Fonetik Değişkenlikler: Türkmencede kalın sıradan olan kelimelere kalın sıradan “-yar”, ince sıradan olan kelimelere “-yär” şimdiki zaman eki getirilir. Yani Türkmencede ek kalınlık incelik uyumuna girdiğinden ve başındaki “y” sesi de kendisinden önce gelen ünlüyü değiştirmede fonetik değişkenlik çok kolay tanımlanacak şekildedir.

Ancak Türkmen ağızlarında kullanılan “-ya/-yâ” şimdiki zaman ekine yazı dilinde de rastlanmaktadır. Bu ek, yazı dilinde kullanılan “-yar/-yär” ekinin sonundaki “r”nin düşmesiyle ortaya çıkmıştır. Çeviri yazılımında bu fonetik değişkenliğin dikkate alınması faydalı olacaktır.

Özel Durumlar: İki lehçeyi karşılaştırdığımızda fonetik değişkenlik dışında soru biçimlerinde dizilişte farklı durumlar bulunmaktadır. Türkmencede soru eki, şahıs eklerinden sonra gelirken; Türkiye Türkçesinde, çokluk 3. şahıs hariç, şahıs ekinden önce gelmektedir. Yukarıda buna ait örnekler verilmişti.

Şimdiki zaman ekiyle örneklenen fonetik değişkenlik konusunun ve ilgili tanımlamaların bütün yapılar da öncelikle çeviri yazılımlarını kodlayacak bilgisayar uzmanlarının kolayca anlayacağı şekilde ayrıntılı olarak ele alınması gerekmektedir. Bu güne kadar sözü edilen çalışmalar yapılmadığı için çeviri yazılımları geliştirilse bile bunların başarı oranı düşük olacaktır.

Bütün bunların yapılmasında şöyle bir sıra izlenebilir: Öncelikle bir lehçenin gramer bakımından kendi iç farklılıkları bütün ayrıntılarıyla ortaya konulmalı, ardından hedef lehçenin kendi içindeki farklı durumları ele alınmalıdır. Daha sonra bu iki çalışma da dikkate alınarak iki lehçenin gramer yapıları eşleştirilmeli, eşleştirilemeyen yerlerde de nasıl bir yol izlenirse çeviri yazılımlarına daha uygun olacağı tespit edilmelidir.

Sağlıklı çeviri yazılımları, yukarıda değindiğimiz uygun sözlüklerin hazırlanması ve çeviriye yatkın karşılaştırmalı gramerlerin yazılmasından sonra geliştirilebilir.

İlk makine çevirisi denemelerinde, basit de olsa, iletişimi kolaylaştıracak düzeyde program veya programların yazılımı gerçekleştirilebilir. Ancak daha gelişmiş çeviri metinlerinin ortaya çıkması, büyük bir ekip oluşturulması ve bu işe bol kaynak ayrılmasından sonra mümkün olacaktır.

Dilbilim Verilerine Göre Türk Lehçeleri Arasında Geliştirilecek Çeviri Yazılımlarının Tahminî Başarı Yüzdeliği:

Başka ortaklıkların yanında ek, kelime, kelime grubu gibi öğelerin ortaklık düzeyi dikkate alındığında iyi geliştirilmiş bir yazılım programının diğer cumhuriyet lehçelerinden Türkiye Türkçesine yapacağı aktarmanın (çevirinin) tahmini başarı yüzdesi şu şekilde olabilir:

Azerice – Türkiye Türkçesi: % 80

Türkmençe – Türkiye Türkçesi: % 75

Özbekçe – Türkiye Türkçesi: % 70

Kırgızca – Türkiye Türkçesi: % 70

Kazakça – Türkiye Türkçesi: % 65

Ortaya konan yazılım, sonradan daha da geliştirilirse, bu oranlar, % 5 ilâ 10 arasında artış gösterecektir.

Öneri: Çeviri yazılımlarıyla ilgili dilbilim çalışmalarının çağdaş Türk lehçeleriyle ilgili yüksek lisans ve doktora tezleri verilerek yaptırılmasını öneriyoruz. Nasıl bir çalışma yapılacağı konusuna, bilgisayar uzmanlarıyla dilbilimciler beraber karar verirlerse, sağlıklı sonuçlar elde edilecek ve önümüzdeki yıllarda başarılı çeviri yazılımları geliştirilebilecektir.

KAYNAKÇA

Basılı Kaynakça

- Aksan, Doğan, 1969, "Türkçe Araştırmalarında Yeni Yollar", Türk Dili Araştırmaları Yıllığı-Belleten 1969, Türk Dil Kurumu Yayınları, Ankara, s. 45-55.
- Aksan, Doğan, 1989, "Kavram Alanı-Kelime Ailesi İlişkileri ve Türk Yazı Dilinin Eskiliği Üzerine", Türk Dili Araştırmaları Yıllığı-Belleten 1971, 2. Baskı, Türk Dil Kurumu Yayınları, Ankara, s. 253-262.
- Cebiroğlu, Gülşen, 2002, Sözlüksüz Köke Ulaşma Yöntemi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü Yayınlanmamış Yüksek Lisans Tezi, İstanbul.
- Cebiroğlu, Gülşen-Adalı, Eşref, 2002, "Sözlüksüz Köke Ulaşma Yöntemi", TBD 19. Bilişim Kurultayı, İstanbul.
- Cebiroğlu, Gülşen-Tantuğ, A. Cüneyd-Adalı, Eşref-Erenler, Yaşar, 2003, "Sentetik Türkçe Sözcük Kökleri Üretimi", International XII. Turkish Symposium on Artificial Intelligence and Neural Networks – TAINN, Çanakkale.
- Coşkun, Volkan, 2000, Özbek Türkçesi Grameri, Türk Dil Kurumu Yayınları, Ankara.
- Ercilasun, Ahmet B. vd., 1991, Karşılaştırmalı Türk Lehçeleri Sözlüğü I (Kılavuz Kitap), Kültür Bakanlığı Yayınları, Ankara.

- Göz, İlyas, 2003, Yazılı Türkçenin Kelime Sıklığı Sözlüğü, Türk Dil Kurumu Yayınları, Ankara.
- Kara, Mehmet, 1997, Ata Atacanov'un Şiirleri I-II (Giriş-Metin-Aktarma-İnceleme-Gramer Dizini), Türk Dil Kurumu Yayınları, Ankara.
- Kara, Mehmet, 2004, Ayır Düşmüş Kelimeler, Çağlar Yayınları, Ankara.
- Kara, Mehmet, 2005, Türkmen Türkçesi Grameri, Gazi Kitabevi, Ankara.
- Kara, Mehmet, 2004, "Çağdaş Türk Lehçeleri Üzerine Bilgisayar Merkezli Çeviri Çalışmaları", TİKA I. Uluslararası Türkoloji Sempozyumu (31 Mayıs-4 Haziran 2004, Kırım-Ukrayna).
- Kara, Mehmet-Karadoğan, Ahmet, 2004, Türkmen Türkçesi-Türkiye Türkçesi Deyimler Sözlüğü, Çağlar Yayınları, Ankara.
- Karadoğan, Ahmet, 2004, Türk Lehçeleri Arasında Aktarma Sorunları-Türkmen Türkçesi-Türkiye Türkçesi Üzerine Bir İnceleme, Kırıkkale Üniversitesi Sosyal Bilimler Enstitüsü Yayınlanmamış Doktora Tezi, Kırıkkale.
- Karakaş, M. Ümit-Hüseynov, Aydın, 2004, "Türkiye Türkçesi ile Azeri, Kazak, Türkmen, Özbek, Kırgız Dilleri Arasında Bilgisayarlı Çeviri İçin Bir Olurluk İncelemesi", Türk Cumhuriyetleri Bilgi Teknolojileri Çalışma Grubu IV. Bilişim İş Birliği Forumu (15-17 Aralık 2004, Bakü-Azerbaycan).
- Kaşgarlı, Sultan Mahmut, 1992, Modern Uygur Türkçesi Grameri, İstanbul.
- Koç, Kenan-Doğan, Oğuz, 2004, Kazak Türkçesi Grameri, Gazi Kitabevi, Ankara.
- Sarıkaya, Yakup, 2003, Berdi Kerbabayev'in Aygıtlı Âdim Romanı (İnceleme-Metin-Sözlük), Gazi Üniversitesi Sosyal Bilimler Enstitüsü Yayınlanmamış Doktora Tezi, Ankara.
- Söygeçov, Mıratgeldi, 1999, Türkmen Dilinin Grammatikası (Morfologiya), "Ruuh", Aşgabat.
- Timurtaş, Faruk K., 1979, Uydurma Olan ve Olmayan Yeni Kelimeler Sözlüğü, İstanbul.
- Tuğlacı, Pars, 1982, Türkçede Anlamdaş ve Karşıt Kelimeler Sözlüğü, İnkılâp ve Aka, İstanbul.
- Uğurlu, Mustafa, 2004, "Türk Lehçeleri Arasında Kelime Eş Değerliliği", Bilig, Bahar 2004, Sayı 29, s. 29-40.
- Vandewalle, Johan, 1992, "Orta Asya Türkçesiyle Türkiye Türkçesinin Arasındaki Farklar Üzerine", Dil Dergisi (Language Journal), Sayı 5, Mayıs 1992, s. 15-20.
- Yaman, Ertuğrul, 2000, Türkiye Türkçesi ve Özbek Türkçesinin Söz Dizimi Bakımından Karşılaştırılması, Türk Dil Kurumu Yayınları, Ankara.
- Zülfikar, Hamza, 1991, Terim Sorunları ve Terim Yapma Yolları, Türk Dil Kurumu Yayınları, Ankara.
- İnternet Kaynakçası
- Altıntaş, Kemal, "Doğal Dil İşleme Çalışmaları ve Türkçe",
(<http://www.teknoturk.org/docking/yazilar/tt000044-yazi.htm>) – 03.05.2005
- Cebiroğlu, Gülşen-Adalı, Eşref, 2004, "An Affix Stripping Morphological Analyzer For Turkish",
(<http://www.cs.itu.edu.tr/~gulsen/pub/411-072.pdf>) – 03.05.2005
- Çakıroğlu, Ünal, "Türkçe İçin Doğal Dil İşleme Çalışmaları-1",
(<http://www.teknoturk.org>) – 03.05.2005
- Oflazer, Kemal, "Two-level Description of Turkish Morphology"

(<http://acl.ldc.upenn.edu/E/E93/E93-1066.pdf>) – 03.05.2005

Taniş, M. Burhan, “Yapay Zeka ve Dilbilimi Üzerine”,

(<http://www.akdenizyazilim.org/default.asp?konu=makaleoku&mno=m1002>) – 03.05.2005

(<http://www.tdk.gov.tr/tdksozluk/sozara.htm>) – 08.05.2005