

***DELIVERY YAPILIR*: SINGLE OCCURRENCE OR CO-SELECTION?**

*Petek Kurtböke*<sup>1</sup>  
*Monash University, Melbourne (Australia)*

**1. Where does *DELIVERY YAPILIR* come from?<sup>2</sup>**

This paper reports on an ongoing study of Turkish-English language contact in Australia where phrases made up of English and Turkish elements are common in Turkish texts. The most illustrative one for the purposes of this paper is *delivery yapılır* (delivery can be arranged). This phrase, like many others, comes from *the Ozturk Corpus*, a collection of 1000 machine-readable texts produced by and for the Turkish migrants in Australia between 1980 and 1995 (for more information on the corpus see Kurtböke, 1996; 1997). The texts that make up the corpus are of two types: newspaper texts and information leaflets.

Firstly, both from a synchronic and a diachronic perspective community newspapers in immigrant settings provide invaluable material for language contact research. In fact, these texts are a rare record<sup>3</sup> of naturally occurring language in immigrant settings across time. In terms of linguistic analysis, however, they remain under-exploited. Such a study, of course, is and has been possible in Australia thanks to the abundance of the Turkish Community newspapers published on a relatively regular basis over the past two decades. The newspaper situation is inevitably different in European countries with considerably larger Turkish communities, due to the factors such as geographical proximity to Turkey, and easily received satellite transmissions from Turkish Radio and TV. Secondly, information leaflets are similarly produced for migrants in the migrant's original language but the provider in this case is the Australian government and various public offices, not the migrants themselves. These texts also remain under-exploited although their nature suits language contact research.

The arguments deriving from the analysis of the *Ozturk Corpus* are also supported with examples from a monolingual corpus, *The Bank of English*. There are two reasons for this. Firstly, behind the compilation of the *Ozturk Corpus*, lies the sound basis provided by two decades of corpus research carried out by John Sinclair and colleagues at the University

---

<sup>1</sup> Linguistics Department, Monash University, Wellington Rd., Clayton, Victoria 3168; E-mail: [nebike.Kurtboeke@arts.monash.edu.au](mailto:nebike.Kurtboeke@arts.monash.edu.au)

<sup>2</sup> This article is based on a paper presented at SIB'97 entitled: "Delivery yapılır: Disappearing L1 or emerging L3?".

<sup>3</sup> Ethnic radio broadcasts could also be used for the purposes of this project however, due to technical problems, SBS Turkish Programs Unit cannot build an archive of past programs (Mr. T. Yenisey, personal communication, 1995).

of Birmingham. Secondly, the lexical operations under investigation show similarities whether the corpus used contains mixed elements or is monolingual.

Traditionally, spoken data have been central in the characterization of language contact phenomena whereas in this study the written mode is central. It also uses computational tools to investigate the regular patterns in written texts which are more difficult to establish manually on the basis of spoken data from individuals. It therefore questions the traditional way of seeing codeswitching as a conversational phenomenon.

## 2. Is DELIVERY a single switch?

The following concordance lines from the *Ozturk Corpus* show the occurrences of *delivery* in the Turkish texts:

yv235 005fb	akşam saat arasında	delivery	yapılır
yv275 001fb	akşam saat arasında	delivery	yapılır
yv275 0011fb	sipariş alınır ücretsiz	delivery	yapılır
yv503 006fb	yere ücretsiz	delivery	yapılır

There are two views on the treatment of lexemes such as *delivery* in language contact data. In one view, there are slots in sentences as in the set above. Let us take *ücretsiz \_\_\_\_\_ yapılır* which translates into English *as free \_\_\_\_\_ can be arranged*. When *delivery* finds itself a slot like this, it jumps in and fills up the slot, and the analysis does not go any further. This is the conventional way of seeing individual loans in data. In the other view, however, the relationship is seen as one of co-selection. If we look at the possibility of *yapılır*'s (can be arranged) selecting *delivery*, it is very high indeed as this verb can select a huge number of nouns (other examples from *Ozturk corpus* *satış* (sale), *nakliyat* (transportation), *tamir* (repair), etc.) and this does not seem particularly significant. However, if we look at the possibility of *delivery*'s selecting *yapmak*, which is very high, the relationship becomes interesting. In other words, *delivery* does not select itself another verb, and even more interestingly in the passive form. Thus, *delivery* and *yapmak* co-select each other and other combinations do not seem likely. When we check the significant collocates of *delivery* in the *Bank of English* we find the following significant ones:

<i>collocate</i>	<i>frequency of occurrence</i>
special	454
home	433
service	381
mail	134
your	282
free	128

express	90
courier	79
arrange	65
food	88

The idea here is that a word form rarely occurs in isolation and when it is used in another language it will inevitably select a number of collocates. The question is “how does this selection occur?”. If we claim that the new collocate will be the translation equivalent of its collocate in the original language we can see that this does not always work. In fact, *delivery yapılır* is a good example of this as *yapmak* translated literally means *do* or *make* and is not an equivalent of *arrange*, the only verb that collocates with *delivery* as shown in the list above. Moreover, there are 65 co-occurrences of this collocation in the *Oznews* component of the *Bank of English*, and out of these only 7 appear in the passive.

Therefore, we need to look for an explanation elsewhere. When we approach the case of *delivery yapılır* from an equivalence perspective (i.e. conventional Language Contact), a frequently asked question is: “The word *delivery* has an equivalent in Turkish, so, why does the Turkish speaker use the English word?”. Here, what the researcher has in mind is the dictionary equivalent of course. Two objections arise. Firstly, recent corpus work on bilingual dictionaries shows that the lexicographer’s intuition which has for centuries played a major role in compilation is misleading. Secondly, parallel corpora in different languages show that equivalences should be considered in relation to respective co-texts (see for example *IJL* Special Issue, 1996). To return to the opening question of this section (Is *delivery* a single switch?) it concludes from the foregoing discussion that within the context of this study the answer is in the negative. The following sections elaborate on this point with reference to the role of *nominalization* as a linguistic process and the use of *delexicalized verbs* in language contact data.

### 2.1. Nominalization

According to Halliday & Martin (1993: 2-21), nominalization originally developed as a feature of scientific writing which enabled grammatical metaphors and has gradually become the norm also in other types of discourse. Importantly, it should be analysed as an essential resource in discourse construction rather than an arbitrary feature of language (Halliday & Martin, 1993: 61). The question of whether languages are inherently nominalized or simply used by people that way is still to be answered, yet it seems quite common across languages (see for example Halliday & Martin, 1993: 124-132 on English and Chinese).

Nominalization and the construction of nominal groups offer a rich alternative to verbal groups which contain only one lexical element: the verb itself (Halliday, 1994: 73). In contrast, the head of the nominal groups may be preceded and/or followed by various modifiers and this brings the elasticity users need for richer expression. However, these constructions are inadequately treated in grammar books particularly in relation to the function of certain elements involved (Sinclair, 1991: 91). To quote Sinclair's example, a construction such as 'the enthusiastic collaboration of auctioneers', will be explained in rather simple terms as a clause transformed into a nominal group (through noun derivation from verbs) without paying any attention to the function of *of*. Rather than treating nominal constructions in terms of derivation, we should rely on collocation (Sinclair, 1991: 92). This point is significant also for language contact research as the function words in such constructions tends to be ignored and/or interpreted as rare switches. Nominalization is often facilitated by the use of delexicalized verbs which seem to play a significant role in language contact.

## **2.2. The role of delexicalized verbs in language contact**

Corpus research has shown that (Sinclair, 1991: 113):

“[t]here is a broad general tendency for frequent words, or frequent senses of words, to have less of a clear, and independent meaning than less frequent words or senses. These meanings of frequent words are difficult to identify and explain; and, with the very frequent words, we are reduced to talking about uses rather than meanings. The tendency can be seen as a progressive *delexicalization*<sup>4</sup>, or reduction of the distinctive contribution made by that word to the meaning”.

Corpus evidence reveals the extent to which such frequent words without independent meaning operate in the construction of normal texts (Sinclair & Renouf, 1988), and it is unhelpful to attempt to analyse them grammatically (Sinclair, 1991: 113).

*Delexicalization* has not been studied extensively as a linguistic phenomenon, and limited literature is available on a limited number of delexicalized nouns and prepositions (e.g. Sinclair, 1989), intensifiers (Partington, 1993), and verbs and adjectives (Sinclair and Renouf, 1988; Sinclair et al., 1996). The notion of delexicalization is central to the study of collocation as the delexicalized word acts in conjunction with other words and shares their meaning (Sinclair & Renouf, 1988; Partington, 1993). Typically, *have*, *take*, *give*, *do* and *make* can be listed among the most delexicalized verbs in English (Collins Cobuild English Grammar, 1990: 147).

“The primary function of *make*, for example, is to carry nouns like *decision/s* [...] thereby offering the alternative phraseology ‘make your own decisions’ to ‘decide on something’ [...] and so on. Which of the two formulations to choose is obviously a strategic matter in text creation, but the delexical option is firmly there” (Sinclair and Renouf, 1988).

Traditionally, three verbs in Turkish, *yapmak*, *etmek*, *olmak* have received similar treatment in the grammars and dictionaries of Turkish. While it has been recognised that these verbs hold some sort of special status in the Turkish lexicon, it is impossible to find a satisfactory description in the standard grammars and dictionaries of Turkish with reference to their delexicalized function particularly with reference to nominalization (Kurtböke, 1998). They often refer to *etmek* as the ‘compound verb’ or ‘auxiliary verb’. It is specified that most loan words from Arabic and Persian operate in Turkish as nouns in *noun + etmek* construction, regardless of their grammatical class in the original language; and there is hardly any mention of *yapmak* with reference to its delexical function (or any other verb for that matter). For example, Underhill’s grammar (1976: 246) touches briefly on this issue:

“Many verbs in Turkish are compounds, formed by a noun indicating an action, followed by the auxiliary verb *etmek*. The noun is usually, although not always, of Arabic, Persian, or European origin [...]. Thus, with the word *telefon* ‘telephone’ we get the verb *telefon etmek* ‘to telephone’; the noun remains invariable, and the verb is conjugated as any other verb”.

Similarly, we find such explanations as the following in *The Concise Oxford Turkish-English/English-Turkish Dictionary* (1959: 98):

“ETMEK is the verb most commonly used to make a composite verb, chiefly with Arabic nouns, e.g. *zannetmek*, to think; *sarfetmek*, to spend; when the noun is of two syllables, the noun and the verb are usually written separately, e.g. *hizmet etmek*, to serve; *telefon etmek*, to telephone”.

The missing information in such descriptions is the frequent nouns or nominal groups which co-occur with these verbs and carry most of the meaning. Whether delexicalization is a universal feature of all languages has yet to be demonstrated however ongoing projects involving parallel corpora show that it is common across the language pairs under examination (Pérez-Hernández, 1996). Romaine (1995) examines a number of similar verbs in Panjabi, which she typically refers to as compound verb constructions. These constructions are formed by a verb, noun or adjective and an ‘operator’. “The operators comprise a small class of simplex verbs with lexical meaning in their own right. The main ones are *krna* ‘to do’ and *hona* ‘to be/become’. The basic meaning of the compound is determined by the first

---

<sup>4</sup> Emphasis added.

element and modified by the verbal operator” (1995: 131). Similarly, Pfaff (1990) gives many examples of such constructions from her Turkish-German data and following Romaine she calls these verbs operators: “A particularly striking difference between the Group A and Group B children is seen in the incorporation of German verbs into Turkish, most typically as German infinitive + Turkish auxiliary *yapmak*”. She lists the following operators which appear in her data frequently: *etmek*, *olmak*, *atmak*, *vermek*, *almak*, *görmek* and *yapmak*. Various other ‘compound verb’ constructions have also been reported elsewhere (e.g. Kachru, 1978; Appel & Muysken, 1987).

If, as argued in the previous section, languages have a tendency to nominalize, delexicalized verbs will then be among the most frequently used items whether the data in hand contain mixed or monolingual constructions as their major function seems to nominalize other verbs.

### **3. A lexical approach to language contact**

The points made by Sinclair et al. above with respect to nominalization and delexicalization naturally derive from the *Bank of English*. The undeniable relevance of these two lexical processes to the study of language contact strongly suggests the suitability of a lexical approach to language contact data. Briefly, a lexical approach to linguistic analysis rejects the view that grammar and lexis are distinct systems. Its main argument is that “each item or structure is used in its own right, rather than as compared with members of the same or a contrasting paradigm” (Francis, 1993). Traditional grammars have not considered the role of lexis in their search for structural restrictions and “in many cases there is no mention at all of the fact that a structure may be lexically restricted other than by considerations of word-class and related factors” (Francis, 1993; see also Stubbs, 1996: 40, for the main arguments of this approach).

This approach looks particularly promising with respect to an unresolved problem in language contact, the boundaries of the lexicon in the contact situation. There are again two views on this. In one view the new lexicon can be established on the basis of the participating languages, L1 and L2. This view, however, does not adequately account for the cases which do not fit within the L1 or L2 lexicons. In the other view, there is a new code developing, L3. The answer again comes from corpus research (Sinclair, 1996b):

“To build an adequate lexicon, we must start with usage. As speakers of the language, even as experts in its lexical structure, we cannot reliably anticipate usage, and so we have to study large samples of the language to uncover the regular patterns”.

### 3.1. Empty L3 lexicon first

Thus, a major shift is required in the way the lexicon is perceived. Sinclair (1996b) further explains how we can achieve this:

“The lexicon is considered empty at the start because nothing appears in it except what is gleaned from the study of the language in use –nowadays, through the study of corpora. There is no assumption that meaning attaches only to the word; it is anticipated that meanings also arise from the loose and varying co-occurrences of several words, not necessarily next to each other. It is, thus, not possible to compile a list of entries in advance of analysing and interpreting the evidence, because the lexical items are not always words, and each word may enter into a variety of relationships with others to realise lexical items”.

Therefore, if we examine loans in relation to their environment, rather than as singly-occurring items, and observe their co-selection tendencies, we can see that the meaning of a word and its immediate new context have become inseparable and it may gradually acquire idiom status (for the format of such a lexicon see Sinclair, 1996b).

Another major question for those who subscribe to the notion of an emerging mixed code in immigrant settings is *when does the mixed code establish itself?* From the discussion so far the answer appears that the mixed code starts establishing itself when mixed collocations develop a life of their own. In other words, when *the idiom principle* (Sinclair, 1991) starts operating. Another relevant question is that if a noun collocates with only a set of words out of a number of possibilities in L1, does it collocate with a similar set when it is borrowed into L2? If it does, we can assume that the collocational patterns of either of the participating languages are not disturbed in the case of contact. But, if it starts co-occurring with a different set, is this to be *considered collocational failure*<sup>5</sup> or the starting point of a new code? The answer to this question lies in the observation of usage and the emerging patterns of co-occurrence (Kurtböke, in progress).

### Bibliography

- Alderson, A.D., F. Iz (1959), *The concise Oxford Turkish-English/English-Turkish dictionary*, Oxford, Clarendon Press.
- Appel, R., P. Muysken (1987), *Language contact and bilingualism*, London, Edward Arnold.
- Baker, M., G. Francis, E. Tognini-Bonelli (eds.) (1993), *Text and technology. In honour of John Sinclair*, Amsterdam/Philadelphia, John Benjamins.
- Francis, G. (1993), “A corpus-driven approach to grammar”, in M. Baker, G. Francis, E. Tognini-Bonelli (eds.), *Text and technology. In honour of John Sinclair*, Amsterdam/Philadelphia, John Benjamins, 138-56.

---

<sup>5</sup> Charles Owen’s term.

- Halliday, M.A.K. (1994), *Spoken and written language*, Melbourne, Deakin University.
- Halliday, M.A.K., J. Martin (1993), *Writing science. Literacy and discursive power*, London, The Falmer Press.
- Kachru, B. (1978), "Toward structuring code-mixing: an Indian perspective", *International Journal of the Sociology of Language* 16, 27-47.
- Kurtböke, P. (1996), "Towards a Corpus of Australian Turkish", *University of Melbourne Working Papers in Linguistics and Applied Linguistics* 16, 51-158.
- \_\_\_ (1997), "A corpus-based study of Turkish community newspapers in Australia: A progress report", *Proceedings of the VIIIth International Conference on Turkish Linguistics*, Ankara, Ankara Üniversitesi, 221-30.
- \_\_\_ (in progress), *Migrants mediation mixed code. A corpus-driven study of Turkish-English language contact in Australia*, Ph.D Thesis, Melbourne, Monash University.
- \_\_\_ (forthcoming), "Non-equivalence of delexicalised verbs in bilingual dictionaries", Euralex 1998.
- Partington (1993), "Corpus evidence of language change - The case of the intensifier", in M. Baker, G. Francis, E. Tognini-Bonelli (eds), *Text and technology. In honour of John Sinclair*, Amsterdam/Philadelphia, John Benjamins, 177-92.
- Pérez-Hernández (1996), "A pilot study on translation equivalence between English and Spanish", in J. Sinclair, J. Payne, C. Pérez Hernández (eds.), *Corpus to corpus: A study of translation equivalence*, *IJL Special Issue*, 218-37.
- Pfaff, C. (1990), "Mixing and linguistic convergence in migrant speech communities: linguistic constraints, social conditions and models of acquisition", *Papers for the Workshop on constraints, conditions and models*, ESF Network on Code-Switching and Language Contact, 120-53.
- Romaine, S. (1995), *Bilingualism* (2<sup>nd</sup> edition), Oxford, Blackwell.
- Sinclair, J. (1989), "Uncommonly common words", in M.L. Tikoo (ed.), *Learners dictionaries: The state of the art*, SEAMEO Regional Language Centre, 135-52.
- \_\_\_ (1991), *Corpus concordance collocation*, Oxford, Oxford University Press.
- \_\_\_ (1996a), "An international project in multilingual lexicography", in J. Sinclair et al. (eds.), 179-96.
- \_\_\_ (1996b), "The empty lexicon", *International Journal of Corpus Linguistics* 1(1), 99-120.
- Sinclair, J., A. Renouf (1988), "A Lexical syllabus for language learning", in R. Carter, M. McCarthy (eds), *Vocabulary and Language Teaching*, London, Longman, 140-60.
- Sinclair, J., J. Payne, C. Pérez Hernández (eds.) (1990), *Collins cobuild English grammar*, London, Collins.
- \_\_\_ (1996), *Corpus to corpus: A study of translation equivalence*, *IJL Special Issue*.
- Stubbs, M. (1996), *Text and corpus analysis*, Oxford, Blackwell.
- Underhill, R. (1976), *Turkish grammar*, Cambridge, Mass., The MIT Press.