Vladimir Ivir

Lexical Gaps: A Contrastive View

0   Human language functions as a system at all levels at
which it is customarily analyzed — i.e., phonological, gramma-
tical and lexical. Though the linguist's analytical tools are not
equally adequate for describing the system at each of the
three levels, he has to recognize the systemic nature of his
material and to look for ways in which it can best be descrip-
tively organized. In the case of the lexical level, one approach
that has been tried has been through the concept of lexical
(or semantic, as used by some authors) fields Different
theories of semantic fields have been formulated (cf. Geckeler,
1971 for a survey of authors and their views), departing more
or less from the original — and not always unambiguous —
formulations of Trier (1931, 1934). However, all the versions
of the semantic field theory view the vocabulary of a language
as a system and propose to analyze and describe it in terms
of its components, which are groups of inter-related words
held together by virtue of their belonging to one particular
conceptual area (or field). Thus, they reflect lexically those
properties of a particular segment of the external world which
the speakers of the language in question have found important
enough to have to talk about them. Under these circumstances
it is clear that words belonging to a given field will be semanti-
cally inter-related so that the sense of each individual word
will be dependent on the senses of all neighbouring words in
the same field. (For the use of the term 'sense', and the way
sense is distinguished from reference, cf. Lyons, 1977:197ff.)

1   Directly related to the systemic view of lexis is the ques-
tion of gaps in the system: is the lexical system of a language
complete and fully integrated or is it incomplete, with holes
at certain points? Furthermore, are lexical fields as subsystems

complete or do they display gaps? The question has been raised repeatedly and answered both affirmatively and negatively. Trier himself (1934:429, quoted by Geckeler, 1971:136) firmly rejected the idea that there may be gaps in the system: "Das Seinsbild einer Sprache is ein Kontinuum, es enthält keine Lücken und blinden Flecke für den Sprachgenossen". Most other authors recognize the existence of lexical gaps as a phenomenon parallel to gaps in the phonological and grammatical parts of the system of language. Kandler (1959) has the word *Lücke* in the title of his paper, Marouzeau (1963:198—212) devotes a whole chapter to *lacunes* in French, Geckeler (1971:134—144) and Lyons (1977:301—305) have sections on gaps, and Lehrer (1970, 1974) devotes an important part of her work to lexical gaps.

Part of the problem in trying to answer the question of the existence or otherwise of gaps in the lexical system of language is definitional. If every language serves perfectly the needs of its native speakers and enables them to express everything that they need to express — as it does, because that is how it has evolved — then Trier is right and no gaps exist in the lexicon. At least, native users of a language are not aware of any as long as they remain within the world in which, and for which, that language has been developed. Its vocabulary as a whole, and the lexical fields into which it is intuitively subdivided, function in a systematic fashion, with every item in a field being dependent for its sense on the neighbouring items and with every field being defined in interaction with other neighbouring fields. When new, previously unlexicalized, aspects of the external words need to be lexicalized, the language does it naturally and that particular part of the lexical system changes to become a new system (in terms of the general systems theory) — again complete and self-contained, just as the previous system was. Native speakers do not feel that a gap has been filled, or perhaps that another gap still remains to be filled.

Another way in which the notion of a lexical gap can be understood is the way in which a linguist studying a lexical system (but not a native speaker) understands it. The linguist's system should presumably agree with the native speaker's intuitions and be psychologically sufficiently real so that a linguistic description can claim to be an explicit and "correct" account of what the native speaker intuitively knows about his language. When lexical fields are set up analytically and examined as systems, then gaps become immediately apparent. They are accidental 'matrix gaps' which show up when "related lexical items are analysed into semantic features and

placed on a chart or matrix" (Lehrer, 1974:97). Understood in this way, lexical gaps no not include every instance of the failure to lexicalize aspects of the external world. In particular, they do not include instances in which a language fails to lexicalize semantic structures contradictory to the native speakers' perceptual experience of the external world or those not available in their culture. Thus, the absence of lexemes for 'male mare' or 'candle-operated TV set' does not mean, in this sense, gaps in the system, since there is no system of male female animals or candle-operated electronic equipment. On the other hand, since English does distinguish lexically between male and female poultry and since it also does lexicalize the semantic feature 'young of' in the lexical field of poultry, its failure to do so in the case of turkey is an example of a gap in the system:

Semantic feature	male	female	'young of'
Type of bird	cock/rooster gander drake turkey	hen goose duck —	chicken gosling duckling —

In this example, three out of the four lexical items have undergone the same lexicalization process and the pattern may be said to be quite clear. In other cases, however, only one item may have had a certain lexicalization rule applied to it and it is not immediately obvious that this is the system and that the failure of other items to lexicalize in the same way gives rise to gaps in the system. To take an example given by Grzegorek (1977:16), in the semantic field of sensory verbs (including *look, listen, smell, taste*) only the verb *look* undergoes lexicalization for the semantic feature 'with special attention' to give the verb *scrutinize*. The other three verbs do not have lexemes to correspond to *scrutinize* and it is doubtful whether this should be interpreted as a case of gaps in the system.

2 The discussion of semantic gaps has so far concerned only one language at a time and gaps that were established were intra-language (i.e., intra-systemic) gaps. Contrastive analysis of pairs of languages (or more languages) may very well proceed by placing the semantic fields of individual languages against one another, observing how they are matched, noting the gaps in one where another has a lexical item,

and writing a differential calculus which could be expressed either verbally or formulaically (cf. Grzegorek, 1977:16—19). The *tertium comparationis* in this approach is the semantic field, which is defined with reference to certain aspects of the external world. Yet, in spite of this fact, the danger of circularity is not avoided. The external world is structured to a certain extent, and structures which are perceived may serve as elements in defining semantic fields. But a good deal of structuring is done by language itself. The dialectic of the human linguistic activity is such that language reflects the world around us, but — once there — it is part of that world and to that extent helps to shape it for us. For that reason the semantic fields which are set up for a given language are peculiar to that language and are not readily matched by semantic fields of another language. When fields of two languages are matched in this way, they may not be relatable to structures perceived by human sensory organs but to structures imposed upon the external world by the first language. In this case, the problem of field boundaries, and therefore of field membership for lexical items, becomes a limiting factor for the description. In the example quoted above, for instance, the gaps in the field of sensory verbs appear different if activities denoted by verbs like *sniff* (at something in order to examine it olfactorily), *snoop, pry,* and *prick up one's ears* are also considered; and the relation of *scrutinize* to *look* is not unaffected by the inclusion of *gaze* and *stare* in the same semantic field. The important thing to note is that while the semantic field has been derived from nature (that is, the experience of sensory activities by humans), its shape and boundaries are determined in our case by the English language. And gaps which are established depend partly on what the analyst has decided to include in, or leave out of, a given field.

More importantly, from a contrastive point of view, whole semantic fields — being language-specific — will not be matchable between languages. For instance, it is possible to establish a field of things, buildings, circumstances, etc. which make it easy or possible to do things and which are labelled 'facilities' in English (as distinct from, say, 'devices', or 'structures'). They include buildings (libraries, laboratories as facilities for study; gymnasia, svimming pools, running tracks, stadium as sports facilities; factories, plants, workshops as production facilities; hotels, restaurants as catering facilities), vehicles (buses, trains as facilities for travel), installations (stoves as cooking facilities, bathroom fixtures as washing facilities), pieces of furniture and equipment (beds as sleeping facilities,

projectors as projection facilities), etc. What holds all these otherwise disparate concepts together is their facilitating function. But that function has been brought to prominence by the fact that the English language has a convenient higher term to subsume all the words in that field and to keep them together. Speakers of a language which does not have a superordinate term to focus on the facilitating function, for instance Croatian, would not intuitively relate these words, nor would an analyst working from that language alone necessarily think of grouping them together. However, he would group together those of the concepts discussed here which the Croatian language readily subsumes under *objekti*, i.e., buildings and sections of buildings constructed for a particular purpose: educational buildings as *školski objekti;* any structures built for sport as *sportski objekti;* structures intended for industrial production as *industrijski objekti;* those for catering as *ugostiteljski objekti;* those for vehicular traffic as *prometni objekti* (including roads, bridges, airports, railway lines, but excluding vehicles, with the significant exception of ships, boats, barges, etc., which are covered by the term *plovni objekti*); those for military purposes as *vojni objekti.* In general, it can be said that *objekt* is a generic term used whenever specific reference is difficult to make or is consciously avoided. Its sense is no more than 'something built' and it only gets its full semantic significance in combination with the modifying adjective which specifies its purpose. Of course it remains generic even then, but the genericness is of a much narrower range, as can be seen in the following examples:

> *Tamo stoji jedan objekt.* ('A building/structure stands there.')
> *To je vjerski objekt.* ('It is a religious building.')
> *To je crkva, kapelica, džamija, hram, samostan . . .*
> ('It is a church, chapel, mosque, temple, monastrey . . .')

In the first sentence, reference may have been to any kind of structure serving any purpose whatsoever; in the second sentence, only one of the buildings given in the third sentence may have been referred to.

Looking for a semantic field in Croatian that could correspond to the field of 'facilities' in English, one does not find any. What one does find instead in this case are matching individual items in unmatched fields and — more significantly — differently organized fields, established on the basis of inter-language differences, are important pointers to intra-language gaps. Facilities do come in different clusters (buildings, things, circumstances), but English does not reflect this, just as Croatian does not reflect the fact that a higher, more

abstract, grouping is possible if various objects are seen in terms of their ability to make it possible to do things. Such contrastively established gaps are places in the system at which new developments can be expected (through word formation, semantic adaptation of existing words, direct translation, or borrowing). The gaps in question are conceptual gaps, discoverable not from within a lexical system, but from the advancement of our perception of the external world which may be due either to our physical experience of that world or to linguistic information about other people's perception of it. This is the value of contrastive analysis which far transcends its immediate usefulness as a foreign language teaching tool. Its general linguistic value comes from the fact that each time that it contrasts a language with another language it puts our understanding of how that first language functions into a new perspective — different from the perspective we get when we study it in isolation or when we contrast it with any other language (cf. Filipović, forthcoming).

3   The concept of gaps plays a crucial role in contrastive analysis which aims to show how different linguistic systems relate to one another not merely as abstract constructs but as different expressions of features of the external world. In the case of lexical systems, it is a well-known fact that different languages lexicalize different semantic characteristics and that for that reason inter-language gaps are found at all those points at which a given language fails to lexicalize a characteristic which another language has lexicalized. The failure to lexicalize may be due to the fact that a particular object or phenomenon is not available in the world in which the speakers of that first language live. Gaps of this kind are random and are filled randomly as the need arises (cf. Ivir, 1973). It is no wonder, for instance, that English should have no lexical item to correspond to the Croatian *štrukli* ('cooked pastry filled with cottage cheese'), just as it is no wonder that Croatian should have no word for *Welsh rarebit*.

The second type of gaps are the product of different perceptions of the world by speakers of different languages. Each community perceives linguistically what is important for it culturally, and it makes those distinctions which it needs to ensure adequate communication among its members. Clearly, family relations are taxonomically the same among English-speaking people as they are among Croatian-speaking people. But English speakers have evolved a culture in which it was not important to distinguish, say, an uncle who is one's father's brother (*stric*) from an uncle who is one's mother's

172

brother (*ujak*), or a male cousin (*bratić*) from a female cousin (*sestrična*). These gaps differ from those of the first kind in that they are not true lexical gaps; rather, they are conceptual gaps since the concepts 'father's/mother's side' or 'male/female' are not expressed in English as they are in Croatian. Notice in this connection that the psychological reality of these gaps depends on language contrasts: native speakers of English become aware of them only when called upon to translate from or into Croatian, their problems being of course different in each direction; before that time they are just as unaware of them as are Croatian speakers of the presumably intra-language gaps in *bratić/sestrična* and *djed/baka* ('grandfather/grandmother'), where the distinction 'father's side' *vs.* 'mother's side' is not lexicalized.

It is the gaps of the second type just discussed that are important and interesting. The first type, those that stand for objects or phenomena which do not exist in a given culture, are culturally important but of no more than marginal linguistic interest. The gaps established intra-lingually, valuable though they are as a form of exercise in systematic description, cannot tell us in what ways the world can be linguistically organized and in what ways our own language leaves gaps in the conceptual framework that other languages may have filled and that it too might fill when the need arises.

It should be stressed, however, that at each stage of its development, regardless of the number of gaps, every language functions as a system and satisfactorily serves the needs of its speakers. The lexical items which it has are structured semantically in such a way, according to language-internal rules, that no gaps remain for its speakers within their own culture. English has one word for 'becoming husband or wife' — i.e., *marry;* Croatian have two: *oženiti se* for 'become a husband' and *udati se* for 'become a wife'. (Both languages have one word for 'enter a marriage': English *marry* and Croatian *vjenčati se.*) This is not to say that the English word has two meanings (neither language, for instance, makes that distinction in *divorce* and *razvesti se*, and yet it does not occur to anyone to speak of two meanings in these words — one for ceasing to be a husband and the other for ceasing to be a wife) but rather that in using that part of the Croatian lexicon, or in translating, the native speaker of English will have to make distinctions that his mother tongue has not equipped him to make. It is then that lexical gaps become psychologically, often painfully, real to him. Faced gaps in the lexicon, the translator resorts to compromise which results in under-

translation, overtranslation, or both, hoping that communicative equivalence has nevertheless been achieved.

Contrastive analysis of the lexicon begins not with semantic fields defined in advance for one and the other language but with individual lexical items. We have seen earlier that *facilities* is a lexical item in English which stands for a concept that is not lexicalized in Croatian. But it does have regular and systematic correspondents in that language based on translational equivalence. Thus, *objekti* corresponds to *facilities* when reference is made to building, *sredstva* corresponds to it when reference is made to means or instruments (*travel facilities* are *prometna sredstva*), while *mogućnosti* is a correspondent of a more general meaning which is used when reference is vague (*cooking facilities* are *mogućnosti za kuhanje, sleeping facilities* are *mogućnosti za spavanje,* etc.). At this point we have, in our example, three words of Croatian which belong together — not from the point of view of the Croatian lexical system, but from the point of view of their contrastive relationship to English. A fourth word, *pogodnosti* 'conveniences', also belongs here, and these four words together contain all the semantic features that *facilities* contains, so that no gap remains. In actual language use, however, only one of them is appropriate in a given context of situation because each contributes its own components of meaning, focusing on one part of the semantic field of facilities and leaving others unrepresented as gaps: *proizvodni objekti* are *production plants, proizvodne mogućnosti* — *production possibilities* (including not only facilities but also natural resources, climatic conditions, skilled labour, ready markets, etc.) and *proizvodne pogodnosti* are *production conveniences* (that is, every type of comparative advantage that makes it convenient for one to engage in industrial production).

Turning in the opposite direction, from Croatian to English, and taking the word *objekt* as a point of departure, the contrastive analyst first notes that its sense when used as a correspondent of *facilities* is a 'man-made structure, building' and that it thus covers only part of the conceptual area covered by the English word. Next, examining the field of application of the Croatian word, the analyst notes that it extends from already built structures to those still under construction or in the various planning stages, from fixed to moving structures, and finally from those which serve their purpose as structures to those which merely support other installations. In English the conceptual focus shifts each time: *ugostiteljski objekt* is *catering facility* or *catering establishment*, and the focus is away from the building; *građevinski objekt* is *con-*

struction project; industrijski objekt is *industrial plant* when completed and *industrial project* while still under construction; *plovni objekt* is *sailing vessel;* and *vojni objekt* is *military installation.*

The same kind of analysis for the other three Croatian words would demonstrate again that each can be used for some kinds of facilities but not for others, that all four of them together cover the entire field of facilities but that each of them individually leaves unlexicalized gaps, and that each shifts the conceptual focus to some other aspect of the world it represents.

4 To summarize, this paper has tried to make several points. The first is that lexical gaps are most profitably seen as conceptual gaps. Viewing them as systemic gaps is fraught with al kinds of difficulties, not the least of which is the indeterminacy of lexical fields within which they are established and the absence of any psychological base for them in the intuitions of native speakers. The second point that has been made is that lexical gaps are inter-language, rather than intra-language, phenomena and that they are connected with the different ways in which different languages reflect their speakers' experience of the external world. The third point is that a contrastive analysis of the lexicon — illustrated here with some English and Croatian examples — reveals both the true nature of lexical gaps and the intricate web of semantic relations among a set of words of one language and between them and a corresponding set of words of another language.

REFERENCES

Filipović, R. (forthcoming), "The Value of Contrastive Analysis to General Linguistics"). A paper read at the XIIth International Congress of Linguists, Vienna 1977. (To appear in the *Proceedings of the Congress.*)

Geckeler, H. (1971), *Strukturelle Semantik und Wortfeldtheorie,* München: Wilhelm Fink Verlag.

Grzegorek, M. (1977), "A Note on Semantic Representation of Lexical Items and Lexical Gaps", *Papers and Studies in Contrastive Linguistics*, Poznań vol. VII, 5—19.

Ivir, V. (1973), "Linguistic and Extralinguistic Considerations in Translation", *Studia Romanica et Anglica Zagrabiensia*, 33—36, 1972—
—1973, 615—626.

Kandler, G. (1959), "Die 'Lücke' im sprachlichen Weltbild: Zur Synthese von 'Psychologismus' 'Sociologismus' ", in *Sprache — Schlussel zur Welt. Festschrift fur Leo Weisgerber*, Düsseldorf, 256—70.

Lehrer, A. (1970), "Notes on Lexical Gaps", *Journal of Linguistics*, VI, 257—61.

Lehrer, A. (1974), *Semantic Fields and Lexical Structure*, Amsterdam: North-Holland Publishing Co.

Lyons, J. (1977), *Semantics*, vol. 1, Cambridge: Cambridge University Press.

Marouzeau, J. (1963), *Aspects du français*, Paris.

Trier, J. (1931), *Der deutsche Wortschatz im Sinnebezirk des Verstandes: Die Geschichte eines sprachlichen Feldes*, Heidelberg: Winter.

Trier, J. (1934), "Das sprachliche Feld: Eine Auseinandersetzung", *Neue Jahrbucher für Wissenschaft und Jugendbildung*, 10, 428—49.