

An abridged version of this paper appeared in S. Van Daele, A. Housen, F. Kuiken, M. Pierrard & I. Vedder, eds., *Proceedings of the Symposium "Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching"*. Brussels, March 29-30, 2007. Brussels: Royal Belgium Academy of Science and Arts; 153-167.

Paraphrase as a Tool for Achieving Lexical Competence in L2

Jasmina Milićević

Dalhousie University & OLST, Université de Montréal
6135 University Avenue, Halifax, N.S., B3H 4P9, Canada
jmilicev@dal.ca

Abstract

The paper argues that paraphrasing competence, i.e., the capacity to produce paraphrases (= synonymous or near-synonymous sentences) is an essential part of lexical competence and of linguistic competence in general. The rationale for this view is that the ability to establish paraphrastic links between utterances of a language *L*, based on lexical relations holding among lexical units of *L*, is a prerequisite for both idiomaticity (\approx naturalness of produced text) and fluency (\approx ease of speaking). It follows that the acquisition of paraphrasing competence should be one of the main goals of language instruction. This, in turn, entails the necessity of developing appropriate teaching tools. One such tool is presented—a learner's *reformulation dictionary*, intended to foster the acquisition and active use of paraphrastic relations.

Keywords

collocations, learner's dictionary, lexical relations, lexical functions, Meaning-Text lexicology, paraphrase, paraphrasing, paraphrasing rules, semantic derivation

1 Paraphrasing and Lexical Competence

From the perspective of text production, the ultimate goal of language instruction in L2 is to help an L2 learner achieve near-native idiomaticity (\approx naturalness of produced text) and fluency (\approx ease of speaking). Both the naturalness and the ease of speaking come from a good command of *lexical relations* and their active use in order to produce, when needed, a varied output, i.e., to *paraphrase*.

Examples in (1) illustrate two major types of lexical relations: *derivational relations* and *collocational relations*. In (1a) we see the verb STEAL and three of its derivatives—nouns denoting, respectively, the agent, the affected object and the action of stealing;¹ (1b) features the noun REBUKE as a collocation base controlling two collocates—a light (= support) verb and an intensifying adjective.

- (1) a. STEAL ~ THIEF_['the person who steals'] ~ LOOT_['that which gets stolen'] ~ THEFT_['the act of stealing']
b. give_[Collocate-1] a sharp_[Collocate-2] REBUKE_[Collocation base]

Examples in (2) illustrate the use of both types of lexical relations in paraphrasing. In (2a), the verb STEAL is replaced by a semantically equivalent collocation featuring the action noun of the verb and a support verb for this noun, and in (2b) the collocation base REBUKE is replaced by a quasi-synonymous noun, REPRIMAND, with the concomitant selection of appropriate collocates for the new base.

¹ Note that I am talking here about semantic derivation, which may but need not be accompanied by morphological derivation.

- (2) a. *STEAL* ≈ *commit*_[Collocate] a *THEFT*_[Action.noun(STEAL) = THEFT; Collocation base]
 b. *give*_[Collocate-1] a *sharp*_[Collocate-2] *REBUKE*_[Collocation base] ≈
*issue*_[Collocate-1] a *stern*_[Collocate-2] *REPRIMAND*_[QSyn(REBUKE) = REPRIMAND; Collocation base]

The ability to paraphrase is part and parcel of linguistic competence (cf. Milićević 2007: 2 and references herein).

On the one hand, paraphrasing is used as a sort of backup preventing a “crash in text generation.” Speaking is an extremely complex process, laden with pitfalls—in the form of restricted lexical co-occurrence, lexical gaps, defective paradigms, etc.—which can only be avoided if we have at our disposal alternative ways of expressing the same meaning, i.e., some “spare paraphrastic variants” that can supplant a sentence whose construction has gone wrong. Examples in (3) illustrate paraphrasing necessary to get around a lexical gap: the verb *ATTACK*_(V) does not have a collocate expressing the meaning ‘intensely with respect to scope’, but the corresponding deverbal noun, *ATTACK*_(N), does, so the “problematic” verb is replaced by a semantically equivalent construction “deverbal N + support V of N” (this is the same type of substitution as the one seen in (2a)).

[Iordanskaja *et al.* 1991: 307]

- (3) a. *The rebels attacked*_(V) *the city* ??? _____ [*intensely with respect to scope*]*
 b. *The rebels launched*_(support V of N) a *whole-scale*_[intense with respect to scope] *ATTACK*_(deverbal N) *on the city*.

On the other hand, paraphrasing is crucially involved in sophisticated linguistic tasks, such as technical writing, abstracting, reformulation or translation; in fact, we can say that these tasks consist mainly in paraphrasing (cf. paraphrasing as a standard component of exercises in rhetoric/stylistics and the increasingly popular view of translation as inter-linguistic paraphrasing). Examples in (4) and (5) illustrate paraphrasing needed to avoid a tedious lexical repetition: in (4), the offending lexeme is replaced by one of its quasi-synonyms, and in (5) the offending collocate of a lexeme L is replaced by another of L’s collocates expressing the same meaning.

- (4) a. *Light CIGARETTES are as lethal as any other* # *CIGARETTES*.
 b. *Light CIGARETTES are as lethal as any other SMOKES*_{[Syn(CIGARETTE) = colloq. SMOKE]*}
 (5) [*Both sides hope that the DEAL*_[Collocation base] *will be reached*_{[Collocate, Realization Verb]*}.]
 a. *But if no DEAL is* # *reached*, ... ⇒
 b. *But if no DEAL*_[Collocation base] *is struck*_{[Collocate, Realization Verb]*} ...

Given the crucial role of paraphrase in language production,² we can put forward the following requirement:

Language instruction in general, and that of an L2 in particular, should be centered on the acquisition of lexical relations and paraphrastic patterns based on these relations.³

This, in turn, implies the necessity of developing *teaching tools* geared toward paraphrasing. Two types of teaching tools can be envisaged: 1) *manuals and learner’s dictionaries* with a paraphrastic component; 2) *diagnostic tests* measuring learners’ paraphrasing ability and serving as an indicator of their over-all language proficiency/the efficiency of instruction.

² Paraphrase is also instrumental in language comprehension; for the role of paraphrase in L2 comprehension, see, for instance, Hsia 2000.

³ I will leave aside the question of when the instruction should start; obviously, this will be different in L1 and in L2.

So far, with rare exceptions (Daunay 2002, Tremblay 2003, Polguère 2004, Apresjan *et al.* 2007), the role of paraphrase and paraphrasing in language instruction has not even been recognized; paraphrase has been used only marginally in exercises and the aptitude to paraphrase has been but cursorily mentioned in L2 requirements. To the best of my knowledge, the only dictionary that makes use of paraphrasing and explicitly acknowledges its role in language teaching is *Lexique actif du français* (Mel'čuk & Polguère 2007). As for diagnostic tests based on paraphrasing, the only attempt to develop one specifically for language teaching seems to be Windsor (1976); Russo & Pippa (2004) propose a paraphrase-based test for measuring the aptitude to interpret. Clearly, then, there is a need to develop paraphrase-centered teaching tools. In this paper, I will present one such tool—a learner's dictionary focusing on the acquisition and use of lexical and paraphrastic relations, called *reformulation dictionary*.

The rest of the paper is structured as follows: Section 2 offers a general characterization of a reformulation dictionary; Section 3 describes a specific reformulation dictionary intended for intermediate-to-advanced learners of French; Section 4 presents a summary of the research reported in the paper and preliminary conclusions.

2 A Teaching Tool Intended to Boost Lexical Competence— a Reformulation Dictionary

This section presents the general principles guiding the elaboration of a reformulation dictionary and illustrates its intended use. The discussion is anchored in the Explanatory-Combinatorial Lexicology (Mel'čuk *et al.* 1995) of the Meaning-Text linguistic theory, a framework that has served for the construction of a number of production-oriented, formalized dictionaries (*DEC*, Mel'čuk *et al.* 1984-1999; *DiCo*, Polguère 2000; *DiCE*, Alonso-Ramos, 2004; *LAF*, Mel'čuk & Polguère 2007; *Computer-Aided Lexical Acquisition Manual*, Apresjan *et al.* 2007). A reformulation dictionary is an Explanatory-Combinatorial Dictionary with the following particularities: it is a learner's dictionary and it is oriented towards paraphrasing. A general characterization of an EC learner's dictionary can be found in Milićević *et al.* (2006).

A reformulation dictionary is a teaching tool intended to support integrated and enlightened lexical acquisition via paraphrasing. *Integrated* acquisition means 'such that it allows for the learning of a lexical item not in isolation but by situating it within the lexical network of a language,' and *enlightened*—'such that it helps the learner grasp regularities in the lexicon, i.e., recurrent links between lexical items, and extract generalizations about semantic and co-occurrence properties thereof.' Since paraphrasing is instrumental in the reinforcement of lexical links already assimilated as well in the activation of new ones, it is an ideal means for achieving this goal.

A dictionary of this type makes use of descriptive tools developed by theoretical linguistics, appropriately adapted to be as learner-friendly as possible. I have in mind, in particular, *lexical functions* (Wanner 1996) and *paraphrasing rules* (Mel'čuk 1992, Milićević 2007) of the Meaning-Text theory, cross-linguistically valid formalisms for describing lexical and paraphrastic relations (for more on these and their possible adaptation to learners' needs, see Section 3). Thus, lexicon acquisition with the aid of a reformulation dictionary is enlightened also in the sense that it presupposes the assimilation by the learner of some knowledge of

linguistics (concepts and formalisms of the science of language, as opposed to knowledge of language).⁴ The advantages of using linguistic descriptive tools in language teaching are easy to see: they not only help learners adopt more efficient learning strategies but also give more depth to the learning experience, since, due to their universality, they allow for easy cross-linguistic and cross-cultural comparisons; cf. Wierzbicka's semantic primitives, which are intended to facilitate exactly this (see, for instance, Wierzbicka 2003 and Goddard & Wierzbicka, to appear).

In order to support the learning process and gradually lead the learner towards independence in the use of a reformulation dictionary, the latter is integrated into a (computer-supported) learning environment comprising tutorials (on different concepts and formalisms used in the dictionary) and exercises.

A dictionary of this type can be envisaged for any language, be it L1 or L2, and at any level of instruction (except, perhaps, beginning L2). While the type and detail of the information to be included into a given reformulation dictionary will vary based on the language/level of instruction, the general organizing principles just expounded remain valid in all cases.

In more concrete terms, lexicographic information in a reformulation dictionary is chosen, organized, and presented as a function of the task it is supposed to help its users accomplish, namely:

Given a sentence S containing a lexical unit L, paraphrase S by manipulating L, i.e., by using the available lexicographic information about L's meaning and co-occurrence properties.

Let me illustrate the type of paraphrasing intermediate-to-advanced learners (of French) may need to resort to. This will help us understand what type of information must be made available to them in a reformulation dictionary developed for the corresponding level. It is relatively easy to extrapolate from this the type of information required for other levels.

Paraphrasing may be needed either in order to correct an ungrammatical text or in order to improve a basically correct but insufficiently "elegant" text; cf., respectively, sentences (6a)-(6b) and (6c).⁵ In both cases, the learner may be guided by remarks, supplied by the teacher, concerning the type of error or impropriety to be addressed; cf. the grammaticality judgments ("*", "?") and error types (in parentheses) in the examples below.

- (6) a. *Ceci [= le cyclisme] offre aussi une bonne occasion pour *regarder (Quasi-synonym) la nature* 'This [= biking] also offers a good opportunity to watch the nature.'
 b. *Souvent, il m'est arrivé de tomber de ma bicyclette et de ?subir une blessure (Collocation, style)* 'Often it happened that I fell from my bike and suffered an injury'.
 c. *Les bicyclettes ne sont pas aussi rapides que les voitures et les camions, mais la société bénéficie de leur usage (Thematic Progression)* 'Bicycles are not as fast as cars and trucks, but society benefits from their use'.

⁴ Cf. Polguère (2004: 1): Inherent characteristics of the lexicon (its enormous size and the non systematicity of its internal organization), **make teaching the notions that structure it more important than teaching lexical data themselves** [translation and emphasis mine—JM].

⁵ These sentences are taken from a corpus of texts written by intermediate-to-advanced FLS learners, used for the elaboration of the *Dire autrement* reformulation dictionary (see Section 3).

Sentence (6a) contains a nomination error: the verb REGARDER ‘look/watch’ has been selected instead of its (richer) quasi-synonym OBSERVER ‘observe’. The indication of the error type will prompt the learner to consult the definition of REGARDER:

REGARDER_{individu ou animal X ~ Y} ≈ ‘X dirige les yeux vers Y dans le but de voir Y’
 individual or animal X ~ Y ≈ ‘X directs X’s eyes towards Y in order to see Y.’

Upon realizing that the central component of the meaning of REGARDER (underlined) is incompatible with the meaning of NATURE (**diriger les yeux vers la nature* [= Y]), the learner will start looking for a quasi-synonym of REGARDER that does combine with this meaning. In this he will be guided by a list of (quasi-)synonyms in the lexical relations zone of the entry for REGARDER, which contains warnings (identified by the symbol “⊗”) in the form of rudimentary explanations concerning semantic differences between the lexemes in question:

{quasi-synonyms}

observer ‘observe’

⊗ regarder les éléments de Y dans le but de mieux connaître Y
 ‘watch the elements of Y in order to better know Y’

fixer#I.1 ‘set one’s eyes on’

⊗ regarder Y sans détourner les yeux
 ‘watch Y without turning one’s eyes from Y’

dévisager ‘look into one’s face’

⊗ regarder le visage de Y
 ‘watch the face of Y’

Etc.

If the above information is insufficient for choosing the right lexeme, the learner will consult and compare the full-fledged definitions of the (quasi-) synonyms of REGARDER:

OBSERVER_{individu ou animal X ~ Y} ≈ ‘X regarde attentivement les éléments de Y les uns après les autres, dans le but mieux connaître Y.’
 individual or animal X ~ Y ≈ ‘X watches attentively the elements of Y one after the other in order to better know Y’

Etc.

Sentence (6b) is deficient because of the improper use of the collocation *subir une blessure* ‘suffer an injury’. This is a perfectly valid collocation, in which *subir* expresses the lexical-functional meaning ‘start having;’ however, this collocation is used mostly in journalistic style (in the context of reporting on events, such as (armed) conflicts and competitive or team sports, which is not the case here). This is what the learner will discover if he looks up the lexical relations zone of the entry for BLESSURE ‘wound’ (the expressions in curly brackets are lexical functions in ‘light’ encoding: they stand respectively, for Oper₁, IncepOper₁ and Caus₁Oper₁ and are literal renditions of the meanings of those functions; more on this will be said in the next section):

...

{avoir ‘have’}

avoir, porter [ART ~]; *souffrir* [de ART]

{commencer à avoir ‘start having’}

presse subir

{se causer de commencer à avoir ‘cause oneself to start having’}

se faire [de ART] // *se blesser*

...

At this point, the learner needs to select a value of another lexical function of BLESSURE with as close a meaning as possible to the initial one (or else attempt a more radical reformulation). The verb *se faire* ‘make oneself’, expressing the meaning of the lexical function ‘cause oneself to start having’, is a fair choice: it does add the meaning of auto-causation but this is perfectly appropriate for the situation being described (someone falling from a bike and hurting himself in the process). Alternatively, the learner may opt for the verb *se blesser* ‘hurt oneself’, a fused

value of the lexical function in question.⁶ In this way, the following reformulations will be obtained: *Souvent, il m'est arrivé de tomber de ma bicyclette et de me faire une blessure <de me blesser>*.⁷

Sentence (6c), a complex clause made up of two coordinated clauses, is grammatical but communicatively deficient: the sentence is about bikes and it would be in order to have BICYCLETTES 'bicycles' as the Theme not only in the first conjunct but also in the second, where this role has been "usurped" by SOCIÉTÉ 'society'. This may be accomplished by substituting to the verb BÉNÉFICIER '(to) benefit' a construction consisting of the copula ÊTRE '(to) be' and the adjective BÉNÉFIQUE 'beneficial', characterizing the second actant of the verb (a Y from which X benefits is beneficial to X): *Les bicyclettes_{Theme} ne sont pas aussi rapides ... mais elles_{Theme} sont bénéfiques pour la société*.

The corresponding information can be found in the paraphrastic zone of the lexical entry for BÉNÉFICIER, containing instantiations of paraphrasing rules in which this verb can participate, with pointers towards the rules themselves (to be found in the tutorial on paraphrasing).

X bénéficie de Y ~ *X tire des bénéfices de X* [Rule n°. xx]
 ~ *Y offre des bénéfices à X* [Rule n°. xx]
 ~ *Y est bénéfique pour X* [Rule n°. xx]
 etc.

And here is the paraphrasing rule (in 'soft' encoding; see below, Section 3):

X V-e [preposition] Y ≈ Y est Adj(V) preposition X.

In the tutorial on paraphrasing, the learner will find necessary explanations concerning the rule, as well as its instantiations involving other lexemes: *X doute de Y* 'X doubts Y' ~ *Y est douteux pour X* 'Y is dubious to X', *X effraie Y* 'X frightens Y' ~ *Y est effrayant pour X* 'Y is frightening to X', *X peut Y-er* 'X can do Y' ~ *Y est faisable pour X* 'Y is doable for X', etc.

The examples above show that a reformulation dictionary for intermediate-to-advanced level needs to contain fairly sophisticated information concerning the meaning, lexical and syntactic co-occurrence of lexemes. In order to be able to find this information, its users need to have quite a bit of linguistic knowledge. See Appendix I for a list of linguistic concepts that seem necessary for this level of instruction.

To conclude this Section, a remark is in order. As the reader will have understood from the preceding discussion, the term *paraphrasing* is used here in a broader-than-usual sense, since even the reformulation of an ungrammatical sentence (with more or less correct meaning) is considered paraphrasing. This broader view of paraphrasing makes it possible to use essentially the same teaching tools and techniques while working with learners having different proficiency levels; thus, for some learners, paraphrasing will be a means of correcting ungrammatical ways of expressing themselves, while others will use it to improve upon their grammatical, but not fully natural output.

⁶ A fused value of a lexical function expresses together, i.e., as one word, the meaning of the function and that of its keyword.

⁷ As this example shows, not only exact paraphrases are admitted but also the approximate ones, changing in a controlled way the initial meaning; this corresponds well to reality—most of the paraphrases people use are approximate paraphrases (cf., for instance, Milićević 2007: 19ff).

3 *Dire autrement* (Say it differently)—a Reformulation Dictionary for Intermediate-to-advanced FSL Learners

Dire autrement (henceforth *DA*) is an electronic reformulation dictionary for intermediate-to-advanced FSL learners, currently developed at the French Department of Dalhousie University (<http://direautrement.french.dal.ca>). In what follows, I will focus on the organization and presentation of lexicographic information in a reformulation dictionary at the micro-level, i.e., within a lexicographic entry, which describes a single lexical unit (= a lexeme or an idiom). Determination of the lexical stock of the dictionary, as well as its macro-organization, i.e., grouping of lexical units into vocables (= polysemous words) and semantic fields, are left out.⁸ Descriptions of the *DA* dictionary that include these aspects of its organization can be found in Milićević & Hamel (in press) and Hamel & Milićević (in press).

Subsection 3.1 presents a standard *DA* entry and identifies in it the sources of paraphrases (i.e., the information that can be exploited for paraphrasing), described **in turn** in subsection 3.2.

3.1 Standard Lexicographic Entry in the *DA* Dictionary

The zones of a *DA* lexicographic entry, their content and (where applicable) the way it is presented are indicated in *Figure 1*. ‘Soft’ encoding refers to learner-friendly adaptations of classical Meaning-Text formalisms; the information relevant to paraphrasing is in bold face.

1. Identification of the Lexical Unit L (= headword of the article) Citation form of L, its part of speech and its grammatical characteristics (e.g., gender, defective paradigm, etc.)
2. Semantic characterization of L a) L’s Semantic Field(s) b) L’s Semantic Label: taxonomic characteristics of L (e.g., fact, act, entity, artifact, etc.) = its “minimal” paraphrase c) L’s Definition: paraphrastic decomposition of L’s meaning d) L’s connotations
3. Syntactic characterization of L L’s Government Pattern [≈ Subcategorization Frame] Presented in ‘soft’ encoding (to be developed)
4. Characterization of Lexical Relations of L Presented in ‘soft’ encoding—in terms of natural language formulas paraphrasing the meaning of corresponding lexical functions
5. Lexical units Related to L (without being quasi-synonymous with it)
6. Paraphrases involving L Instantiations of paraphrasing rules applicable to L, with pointers towards the rules, presented in ‘soft’ encoding (to be developed)
7. Examples of L’s use
8. List of idioms containing L’s signifier
WARNINGS alerting the user to potential difficulties associated with L’s use (may be found in any of the zones 1-4; the zone 5 is a warning in itself).

Figure 1: Template of a *DA* Entry

A *DA* entry is actually a “hybrid” entry in the sense that it combines types of lexicographic information and formalisms that can be found in other EC dictionaries, while at the same time adding information and formalisms peculiar to it. Thus, like *ECD*, it contains a full-fledged

⁸ Let me just mention that one of the important criteria for the selection of the lexical stock is the “paraphrasing potential” of lexemes, correlated to the complexity of their meaning/richness of their lexical relations.

lexicographic definition of the keyword (sub-zone 2c), written in a somewhat simplified *ECD* style; like *LAF*, it uses ‘soft’ encoding of lexical functions (zone 4); and it is the only EC dictionary that contains a paraphrasing component (zone 6), with ‘soft’ encoding of paraphrasing rules. A *DA* entry and an *ECD* entry for the same lexeme are given in Appendix II.

3.2 Sources of Paraphrases in a *DA* Entry

Sources of paraphrases in a *DA* entry are as follows: 1) semantic label [sub-zone 2b]; 2) definition [sub zone 2c]; 3) Government Pattern [zone 3]; and 4) lexical relations [zone 4]. Sub zone [2c] is a minor source of paraphrases only, i.e., paraphrasing possibilities it offers are rather limited,⁹ while other (sub)zones can be either minor or major sources. In what follows, I will discuss in turn (sub)zones [2c] through [4], leaving aside sub zone [2b]. As we shall see, part of the information relevant to paraphrasing that can be deduced from these sources is explicitly indicated in zone (6).

3.2.1 Lexicographic definition of *L*

Lexicographic definition of *L* allows for semantic paraphrasing—via manipulation of (configurations of) meanings. We can either

- 1) replace *L* by its definition, which is a semantic decomposition of *L*’s meaning in terms of simpler meanings (this is a minor source of paraphrases) or
- 2) use *L*’s definition as a ‘point of entry’ for omitting/adding some meanings or changing their configurations (a major source).

Definition of *L* is actually a rule that establishes the equivalence between *L* (more precisely, *L*’s propositional form—a formula containing *L* and its semantic actants) and *L*’s semantic decomposition; cf.:

Rule^{SEM} 1:

[*L*’s propositional form = definiendum] *X empêche Y de faire Z(Y)* ‘*X* prevents *Y* from doing *Z(Y)*’ ≈
 [*L*’s semantic decomposition = definiens] ‘*X* cause que *Y* ne peut pas faire *Z(Y)*’ ‘*X* causes *Y* not to be able to do *Z(Y)*’

- (7) a. *Ce bruit m’empêche de dormir* ‘This noise prevents me from sleeping’. ≈
 b. *Je ne peux pas dormir à cause de ce bruit* ‘I cannot sleep because of this noise.’

Even though the definition of *L* is an exact paraphrase of *L*’s meaning, it cannot always replace *L* in texts, due to its insufficient idiomaticity. The result of substitution is quite acceptable in some cases (*Mon gendre = Le mari de ma fille enseigne l’anglais* ‘My son-in-law = My daughter’s husband teaches English’), but very often it is not (*Jean s’est réveillé *= a cessé de dormir à 7 heures* ‘*Jean* woke up *= stopped sleeping at 7 o’clock’). This source of paraphrase should thus be used with caution.

The number of semantic decomposition rules is equal to the number of lexemes of the language (roughly, a million). Once completed, the *DA* dictionary will contain 1.000 such rules.

⁹ On semantic labels and their use in Meaning-Text lexicography, see Polguère (2003) and Milićević (1995).

Semantic omissions/additions and meaning configurations' modifications are taken care of by semantic reconstruction rules. A rule of this type establishes an equivalence (or a quasi-equivalence) between two configurations of meanings that cannot be related via definitions; cf.:

Rule^{SEM} 2:

'X fait Y habituellement' 'X does Y habitually' ≈

'Y peut faire Y' 'X can do Y'

(8) a. *Cet enfant parle couramment* 'This child speaks fluently'. ≈

b. *Cet enfant peut parler couramment* 'This child can speak fluently.'

Semantic reconstruction rules constitute the Meaning-Text semantic paraphrasing sub system. Since these are fairly new rules, their number has yet to be assessed (my guess is that it should be around a few hundred). The DA dictionary will use only a small number of these rules; see Appendix I.

Semantic decomposition and reconstruction rules as presented above are already in 'soft' encoding; in their 'hard' version they are written as (quasi-)equivalencies between semantic networks. However, DA definitions are simplified with respect to those of ECD; for an illustration, see Appendix II.

3.2.2 Government Pattern [= GP] of L

The GP of L allows for purely syntactic paraphrasing—by changing the syntactic environment of L without changing L itself. This can be done by exploiting either

- 1) alternative realizations of L's actants (this is a minor source of paraphrases) or
- 2) different modifications of L's GP (a major source); cf., respectively, (9) and (10):

(9) a. *Je vous conseille la prudence <d'être prudents>* 'I advise you caution <to be cautious>'.
b. *Jean éprouve du respect pour <envers, vis-à-vis de> ses parents* 'Jean has respect for <towards, vis-à-vis> his parents'.

The second syntactic actant of CONSEILLER#1.1 '(to) advise' (what is advised) can be realized as a prepositionless noun or a verb in the infinitive introduced by the preposition DE, while the second syntactic actant of RESPECT 'respect_(N)' (the object of respect) can be introduced by either of the three prepositions indicated above.

(10) a. *La presse a critiqué le Gouvernement pour sa décision d'augmenter les taxes* 'The press criticized the Government for its decision to raise taxes'. ≈

b. *La presse a critiqué la décision du Gouvernement d'augmenter les taxes* 'The press criticized the Government's decision to raise taxes'.

In (10a), the verb CRITIQUER '(to) criticize' has a three-actantial realization in (X ~ Y for Z(Y)) and in (10b) a two-actantial one (X ~ Z(Y)); these realizations are quasi-synonymous. To put it simply, one can (more or less) indifferently criticize a person for his action or the action (of the person) itself.

Lexemes close in meaning tend to have similar GPs; thus, other "evaluative" lexemes—BLÂMER '(to) blame', DÉSAPOUVER '(to) disapprove', CONDAMNER '(to) condemn', etc.—have the same number of actants and admit the same GP modifications as CRITIQUER. Obviously, this is a valuable generalization that should be brought to the attention of learners.

3.2.3 Lexical functions [LFs] of L

Lexical functions allow for lexical-syntactic paraphrasing—via synonymic lexical substitutions, possibly leading to modifications of the initial syntactic environment. We can either

- 1) replace an element of the value of an LF(L) by another such element or by a literal realization of the meaning of the LF (these are minor sources of paraphrases) or else
- 2) replace L by L' (or a lexical configuration) with which L is in a lexical-functional relation (a major source).

To illustrate the first case, let us consider the following LFs of the lexeme CONSEIL#1.1 'advice',¹⁰ and their respective values; each LF is indicated in its standard notation and in the 'soft' encoding—in fact, a natural language formula paraphrasing the meaning of the LF (Popovic 2003 and Polguère 2004).

'Hard' encoding (in terms of LFs):	Natural language formula	Value
1. Oper ₁	{donner un C.} 'give A.'	<i>donner</i> [(ART) ~ à N _z]; <i>fournir, offrir</i> [ART ~ à N _z]
2. trop de C. + Labor ₁₂	{donner trop de C.} 'give too much A.'	<i>inonder</i> [N _z de ~s], <i>accabler</i> [N _z de ~s]
3. essayer de Caus ₃ Func ₃	{essayer de recevoir un C.} 'try to receive some A.'	<i>demander</i> [(ART) ~ à N _z], <i>solliciter</i> [ART ~ à N _z]

Figure 2: Three lexical functions of the noun CONSEIL#1.1 'advice' and their values

For each of the LFs, any element of its value can—as a rule—be replaced by any other, since they are synonymous; cf.:

- (11) a. *Jean lui a donné* ≈ *a fourni* ≈ *a offert un CONSEIL fort utile* 'Jean gave <furnished, offered> him some very useful advice.'
 b. *Tu les inondes* ≈ *accables de CONSEILS* 'You flood <burden> them with advice.'
 c. *Il faut demander* ≈ *solliciter son CONSEIL* 'It is necessary to ask for <solicit> his advice.'

An element of the value of a given LF(L) can also be replaced by the natural language formula which encodes the meaning of this LF; cf.:

- (12) a. *Tu les inondes de CONSEILS* 'You flood them with advice.' ≈ *Tu leur donnes trop de CONSEILS* 'You give them too much advice.'
 b. *Il faut lui demander un CONSEIL* 'It is necessary to ask for his advice.' ≈ *Il faut essayer de recevoir son CONSEIL* 'It is necessary to try to receive his advice.'

A natural language formula encoding an LF's meaning is of course less idiomatic than the values of the LF (for a given L); thus, the formulas should be used as the "last resort", if the values themselves are not known or cannot be recalled.

Let me now turn to the second (major) case of lexical-syntactic paraphrasing: lexical-functional equivalencies, modeled by means of paraphrasing rules, such as the following two.

Rule^{LEX-SYNT} I: L_(V) ≈ Oper₁(S₀(L_(V)))—II→S₀(L_(V))

- (13) a. *Jean est tombé* [= L_(V)] 'Jean fell' ≈
 b. *Jean a fait* [= Oper₁(S₀(L))]
une chute [= S₀(L)] 'Jean took a fall'

¹⁰ While CONSEIL1.1 is a countable noun, its English equivalent, ADVICE, is not; cf.: *donner un conseil / give a piece of <some> advice*. The entry for CONSEIL1.1 in a reformulation dictionary developed specifically for Anglophone users should carry this information (in the form of a warning, cf. 3.1: 7).

This rule describes a synonymic substitution—a replacement of a verbal lexeme L by a configuration consisting of the corresponding deverbal noun [S_0] and the support verb [$Oper_1$] that takes this noun as its object and the first syntactic actant of the noun as its subject.

Rule^{LEX-SYNT} 2: $L \approx \text{Anti}(L)\text{—ATTR}\rightarrow\text{_NE}\dots\text{PAS_}$

- (14) a. *Jean était absent* [= L] ‘Jean was absent.’ ≈
 b. *Jean n’était pas présent* [= Anti(L)] ‘Jean was not present’

This is an antonymic substitution—a replacement of a lexeme L by a configuration consisting of an antonym of L and the negation.

There are about hundred rules of this type; they constitute the Meaning-Text lexical-syntactic paraphrasing subsystem. Only one tenth of these rules are used in the DA dictionary (see Appendix I).

A learner-friendly formalism for presenting lexical-syntactic paraphrasing rules has yet to be worked out; here are two suggestions.

- 1) From L’s semantic label (cf. p. 7) one can foresee with reasonable accuracy the LFs applicable to L, as well as their values. Thus, lexical units denoting acts admit the LF $Oper_1$, whose default value with them is *faire* <effectuer> ‘do <perform>’; lexical units denoting emotions also admit this LF, which is commonly realized with them as *éprouver* ‘feel’; etc. This information can be incorporated into paraphrasing rules involving $Oper_1$, such as Rule^{LEX-SYNT} 1, which will then look something like this:

$V_{\text{‘acte’}} [X \sim Y] \approx [N_X] \text{ effectue/fait } N_{\text{‘acte’}} [de N_Y]$

E.g.: *analyser* ‘(to) analyse’ ≈ **effectuer** <**faire**> *une analyse* ‘analysis_(N)’; *réparer* ‘(to) repair’ ≈ **effectuer** <**faire**> *une réparation* ‘repair_(N)’

$V_{\text{‘sentiment’}} [X \sim Y] \approx [N_X] \text{ éprouve/ressent } N_{\text{‘sentiment’}} [pour N_Y]$

E.g.: *respecter* ‘(to) respect’ ≈ **éprouver** *du respect* ‘respect_(N)’; *aimer* ‘(to) love’ ≈ **éprouver** *de l’amour* ‘love_(N)’

$V_{\text{‘information’}} [X \sim Y \text{ à } Z] \approx [N_X] \text{ donne/fournit } N_{\text{‘information’}} [à N_Y]$

E.g.: *conseiller* ‘(to) advise’ ≈ **donner** (un) *conseil* ‘advice’; *informer* ‘(to) inform’ ≈ **donner** <**fournir**> *une information* ‘information’

Etc.

Of course, we could try to rewrite in this way the rules involving other LFs.

The limitations of this technique are obvious: a) it is not completely error-proof since not all LFs and not all values of LFs for lexical units with the same semantic label coincide; b) not all rules can be rewritten in this way (e.g., rules describing inversion of subordination, which are more syntactic in nature); and c) a gain in transparency is set off by a considerable loss in generality, since we need to write several ‘soft’ rules to account for a single ‘hard’ one.

- 2) Alternatively, we could rewrite Rule^{LEX-SYNT} 1 in the following way:

$V_{[X \sim Y]} \approx [N_X] \text{ support } V(N) + N_{\text{‘fact of V-ing’}} + \text{PREPOSITION } [N_Y]$.

See also the rule given p. 6. This notation is more general but less transparent than the one suggested above.

Let us now turn to the paraphrastic zone of a *DA* entry [zone 6]. This zone presents explicitly the information that can be deduced from zones [2c], [3] and [4] (with the exceptions of the information coming from the so-called “minor sources” of paraphrases, cf. p. 8). More specifically, it contains instantiations of paraphrasing rules (applicable to the keyword) of the following types: 1) semantic decomposition and reconstruction rules exploiting L’s definition; 2) syntactic rules exploiting L’s GP modifications; and 3) lexical-syntactic rules exploiting L’s lexical functions. Each rule instantiation is linked to the corresponding rule in the tutorial on paraphrasing. For an illustration of the paraphrastic zone, see the entry of CONSEILLER#1.1 ‘(to) advise’ in Appendix II.

4 Summary and Preliminary Conclusions

The paper has argued that lexical competence stems largely from paraphrasing competence, i.e., the ability to exploit lexical relations in order to reformulate one’s discourse. It has pointed out the necessity of developing teaching tools geared toward paraphrasing and has presented one such tool—a learner’s reformulation dictionary anchored in the Combinatorial-Explanatory lexicology of the Meaning-Text linguistic theory. A specific reformulation dictionary for intermediate-to-advanced learners of French has been described.

Teaching tools intended to facilitate lexical acquisition must make use of concepts and formalisms developed by theoretical lexicology, as only the latter allow for coherent and rigorous structuring of (notoriously complex) lexicographic information. This, in turn, entails the necessity that their users—in particular L2 learners—assimilate those concepts and formalisms. This task, often intimidating for learners, can be facilitated if a learner-friendly version of linguistic descriptive tools is adopted. (However, as our experience so far suggests, *learner-friendly* does not necessarily mean ‘less formalization.’)

Acknowledgments

I am grateful to Margarita Alonso-Ramos, Yves Bourque, Marie-Josée Hamel, Igor Mel’čuk, Muriel Péguret and Alain Polguère for their comments on a pre-final version of this paper.

References

- Alonso Ramos, M. (2004). Elaboración del Diccionario de colocaciones en español y sus aplicaciones. In : P. Bataner and J. de Cesaris (eds.), *De Lexicographia. Actes del I Symposium internacional de Lexicografía*, Barcelona, IULA-Edicions Petició, 149-162.
- Apresjan, Yu., Djačenko, P., Lazurski, A. & Tsinman, L. (to appear). O komp’jutornom učebnike leksiki russkogo jazyka [A Computer-Assisted Manual for the Acquisition of Russian Lexicon]. *Russkij jazyk v naučnom osveščanii*.
- Daunay, B. (2002). *La paraphrase dans l’enseignement du français*. Bern: Peter Lang.
- Goddard, C. & Wierzbicka, A. (to appear). Semantic Primes and Cultural Scripts in Language Learning and Intercultural Communication. In: Sharifian, F. & Palmer, G. (eds.), *Applied Cultural Linguistics: Implications for Second Language Learning and Intercultural Communication*.
- Jordanskaja, L., Kittredge, R. et Polguère, A. (1991). Lexical Selection and Paraphrase in a Meaning-Text Generation Model. In: Paris, C. L., Swartout, W. R. et Mann, W. C., eds. *Natural Language Generation in Artificial Intelligence and Computational Linguistics*. Boston: Kluwer. 293-312.

- Hamel, M.-J. & Milićević, J. (in press). Analyse d'erreurs lexicales d'apprenants du FLS: démarche empirique pour l'élaboration d'un dictionnaire d'apprentissage. *Journal of the Canadian Association for Applied Linguistics*.
- Hsia, S. (2000). Grammaticality judgments, paraphrase and reading comprehension: evidence from European, Latin American, Japanese and Korean ESL learners. sunzi1.lib.hku.hk/hkjo/view/10/1000030.pdf
- Mel'čuk, I. (1992). Paraphrase et lexique: la théorie Sens-Texte et le *Dictionnaire explicatif et combinatoire*. In: Mel'čuk, I., Arbatchewsky-Jumarie, N., Iordanskaja, L., Mantha, S. & Polguère, A. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques III*. Montréal: Presses de l'Université de Montréal; pp. 9-59.
- Mel'čuk, I., Clas, A. & Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve: Duculot.
- Mel'čuk, I. & Polguère, A. (2007). *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques du français*. Louvain-la-Neuve: de Boeck Duculot.
- Milićević, J. (1995). *Étiquette sémantiques dans un dictionnaire de type explicatif et combinatoire* [mémoire de maîtrise]. Département de linguistique et de traduction, Université de Montréal.
- (2007). *La paraphrase. Modélisation de la paraphrase langagière*. Bern: Peter Lang.
- Milićević, J., Alonso-Ramos, M. & Hamel, M.-J. (2006). Un dictionnaire d'apprentissage de type explicatif et combinatoire. Principes et techniques d'élaboration. Communication présentée lors du 30^e colloque de l'Association linguistique des provinces de l'Atlantique, Halifax, 3-4 novembre, 2006.
- Milićević, J. & Hamel, M.-J. (in press). Un dictionnaire de reformulation pour les apprenants du français langue seconde. *Actes du 29^e colloque de l'Association linguistique des provinces de l'Atlantique, Moncton, 4-5 novembre, 2005*.
- Polguère, A. (2002). Une base de données lexicales du français et ses applications possibles en didactique. *Revue de linguistique et de didactique des langues (LIDL)* 21: 75-97.
- (2003). Étiquetage sémantique des lexies dans la base de données DiCo. *Traitement Automatique des Langues (T.A.L.)*, 44/2: 39-68.
- (2004). La paraphrase comme outil pédagogique de modélisation des liens lexicaux. In: Calaque, E. & David, J., réds., *Didactique du lexique : contextes, démarches, supports*. Collection « Savoirs en pratique ». Bruxelles: De Boeck; pp. 115-125.
- Popovic, S. (2003). *Paraphrasage des liens de fonctions lexicales* [mémoire de maîtrise]. Département de linguistique et de traduction, Université de Montréal.
- Russo, M. & Pippa, S. (2004). Aptitude to Interpreting: Preliminary Results of a Testing Methodology Based on Paraphrase. *META*, 49: 2; pp. 409-432.
- Tremblay, O. (2003). *Une approche structurée de l'enseignement du lexique en français langue maternelle basée sur la lexicologie explicative et combinatoire* [mémoire de maîtrise]. Département de linguistique et de traduction, Université de Montréal.
- Wanner, L., ed. (1996). *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam/Philadelphia: Benjamins.
- Wierzbicka, A. (2003). *Cross-Cultural Pragmatics. The Semantics of Human Interaction*. The Hague: Mouton / de Gruyter.
- Windsor, Micheline (1976). L'exercice de paraphrase dans l'enseignement du français. *Audio visual language learning* 14/1: 3-8.

Appendix I

Linguistic Concepts and Formalisms in the DA Dictionary

I. Concepts

actant (semantic, syntactic); circumstantial; collocation; government, government pattern; lexical function; lexical relation (derivational, collocational); lexical unit (lexeme, idiom); paraphrase/paraphrasing, paraphrasing rule (semantic, lexical-syntactic); part of speech, Rheme (communicative mark); semantic decomposition; semantic derivative; semantic predicate; semantic field; semantic label; stylistic label; surface syntactic relation (subject, direct/indirect/oblique object, modifier, determiner); Theme (communicative mark), vocable.

II. Lexical functions

1) Standard LFs

- Paradigmatic LFs

Simple:

Gener, Syn, Anti, Conv_{ijkl}; V₀, S₀, S_i, S_{res}, S_{instr}, A₀, A_i

- Syntagmatic LFs

Simple:

Oper_i, Func_i, Real_i, Fact_i, Magn, Bon, Ver, Mult, Sing

Complex:

the above combined with Incep, Cont, Fin, Caus, Perm, Liqu, Anti (ex., IncepOper_i, CausFunc_i, AntiReal_i, AntiMagn)

2) Non-Standard LFs (ex., essayer de CausFunc₀). Used parsimoniously.

III. Paraphrasing rules

1) Semantic paraphrasing rules

- Decomposition rules. In all cases where the result is sufficiently idiomatic.
- Reconstruction rules. Used parsimoniously.

2) Lexical-syntactic paraphrasing rules

(1) $L \equiv \text{Gener}(L)$

*John underwent an operation [= L] \equiv
John underwent a _surgical intervention_ [= Gener(L)]*

(2) $L \equiv \text{Syn}(L)$

*Where is my bicycle [= L] ? \equiv
Where is my bike [= Syn(L)] ?*

(3) $L \equiv \text{non+Anti}(L)$

*John dislikes [= L] Mary. \equiv
John does not [non] like [= Anti(L)] Mary.*

(4) $L \equiv \text{Conv}_{ijkl}(L)$

*John is Mary's husband [= L]. \equiv
Mary is John's wife [= Conv₂₁(L)].*

(5) $L_{(V)} \equiv V_{\text{supp}}(S_i(L_{(V)})) + S_1$

*John fell [= L_(V)]. \equiv
John took [= Oper₁(S₁(L_(V)))] a fall [= S₀(L_(V))].*

*They attacked [= L_(V)] at dawn. \equiv
Their attack [= S₀(L_(V))] took place [= Func₀(S₀(L_(V)))] at dawn.*

(6) $L_{(V)} \equiv V_{\text{real}}(S_i(L_{(V)})) + S_i$

*John was hospitalized [= L]. \equiv
John was taken [= Labreal₁₂(S_{1oc}(L_(V)))] to the hospital [= S_{1oc}(L_(V))].*

(7) $L_{(V)} \equiv V_{\text{supp/real}}(A_i(L_{(V)})) + A_i$

*John knows [= L_(V)] a lot. \equiv
John is [= Oper₁(A₂(L_(V)))] very knowledgeable [= A₁(L_(V))].*

(8) $//f(L) \equiv f(L) + L$

*What a downpour [= //Magn(L = RAIN)]! \equiv
What a heavy [= Magn(L)] rain [= L]!*

Appendix II

Entry for CONSEILLER.I.1 '(to) advise' in the DA Dictionary and in the ECD

DA

CONSEILLER.I.1, verbe, transitif

Caractérisation sémantique

Champ sém.: communication verbale, acte de parole, aide
 Étiquette sém.: communiquer verbalement
 Définition: individu X ~ à individu Z de faire Y =
 X croit que Z veut savoir#I l'opinion#1 de X sur ce que Z doit faire dans la situation en question.||
 X communiqué à Z que faire action Y ou être en état Y serait dans l'intérêt de Z,
 le but de cette communication étant d'inciter Z à Y.

☒ **Emploi performatif est possible: "Je vous/te conseille [Y]"**

Expression des actants

[qui conseille] X = I = N
 [ce qui est conseillé] Y = II = N, de V-inf, "PROPOSITION", LE_(Pron.neutre)
 [à qui on conseille] Z = III = à N, IL_(clitique.datif)

☒ **Expression de Y est obligatoire**

Jean [X] ~ à Marie [Z] la modération <d'être modérée> [Y]; Il [X] le [Y] lui [Z] ~; Jean [X] ~: Sois modérée! [Y]

☒ **Emploi impersonnel est possible: "Il est conseillé (à Z) de [Y]"**

Relations lexicales

{quasi-synonyme} recommander, suggérer
 ☒ **recommander ≈ conseiller vivement; suggérer ≈ conseiller discrètement**
 {antonyme} déconseiller, "atténué" ne pas conseiller#I.1
 ☒ **'conseiller de ne pas faire Y'**
 {X qui C. habituellement} conseiller_(O)#I.1
 {fait de C. ou ce qui a été C. [=Y]} conseil#I.1
 {intensément} fortement, vivement

☒ **Lexies de sens proche**

PROPOSER#2a

Paraphrases

X ~ à Z de faire Y ≈
 1) X **dit** à Z **qu'il sera dans l'intérêt** de Z de faire Y [Par décomposition]
 2) X **recommande** <**suggère**> à Z de faire Y [Règle n° xx]
 3) X **donne** à Z **le conseil** de faire Y [Règle n° xx]
 4) **Le conseil** de X à Z **est** de faire Y [Règle n° xx]
 5) X **déconseille** à Z de faire Y [Règle n° xx]

X ne ~ pas à Z de faire Y ≈

Exemples

Son médecin lui conseilla le repos. | Nous vous conseillons vivement de faire votre demande en ligne. [Source: Google]

ECD III: 175

I.1 X conseille Y à Z = 'En croyant#II que Z veut savoir#I ce que X pense à propos du comportement que Z doit adopter dans la situation en question,|| X communique à Z que X croit#II qu'il serait dans l'intérêt de Z que Z ou quelqu'un lié à Z fasse l'action (liée à) Y ou qu'il soit dans l'état Y – dans le but d'inciter Z à Y.'

N.B.: Peut être employé performativement.

Régime

X = I	Y = II	Z = III
1. N	1. N 2. de V-inf 3. que PROP-subj 4. "PROP" obligatoire	1. à N

1) C_{2,1}, 2,2: M₁(Y) = Z

2) C_{2,3} : M₁(Y) ≠ Z

C₁ + C₂ + C₃: Le gouvernement a conseillé la modération <le statu quo> à la population [...]

Fonctions lexicales

Syn.: recommander, proposer#2a suggérer
 Anti: déconseiller, ne pas conseiller#I.
 ['communiquer qu'il ne serait pas dans l'intérêt de Z que Z fasse Y']
 S₁^{usuel}: conseiller_(O)#I.1
 S₂: conseil#I.1
 Magn_['inciter']: fortement, vivement
 S₁^{usuel}+le locuteur a une attitude négative envers X
 à cause de ses C. mal appropriés: **vieilli** conseiller

Exemples

Le Sénat a conseillé la prudence aux députés de la Chambre. Jean a conseillé à Nicole de se tenir tranquille. André a conseillé à son ami de réparer cet oubli sans délai. Le policier a fortement conseillé au responsable que les grévistes quittent les lieux.