Katalin Wéber

Coreference in understanding a text (for learners of Hungarian as a foreign language)

Introduction

The major goal of my article is to underline the importance of coreference in tasks of reading comprehension. The focused search for co-referential links within a text may help learners of Hungarian as L2 identify and match content and grammatical elements. In order to cast light on the issue and support my idea I will analyze reading comprehension tasks of ECL exam of Hungarian of level B2 and examiners' correct and wrong answers. I claim that in texts of level B2 where not only explicit but inferred meaning might also be implied the text as a whole of successive sentences conveys a way more complex meaning than the mere sum of its components. In order to be able to reveal the implicit content in a Hungarian text of level B2 coreferential awareness is essential. Moreover, coreferential links must be understood to answer simpler questions based on local bits of information of a text.

Reading competence and tasks on level B2

In Hungary as well as in other European countries language exams in the same way as ECL exam (Hegedűs 2003, Szabó 2011) are aligned to the levels of foreign language competence as described by the Common European Framework of Reference:

ECL levels	CEFR levels
	C2
LEVEL C1	C1
LEVEL B2	B2
LEVEL B1	B1
LEVEL A2	A2
	A1

Level B2 (or "vantage" level) refers to the proficiency of an independent language user who can read with a large degree of independence texts of various kinds and has a broad active vocabulary but may experience some difficulties when reading low frequency idioms. According to the specifications of ECL language exam a level B2 learner can start studies in secondary schools and at colleges and universities of the target language country.

As in each skill of ECL language exam reading comprehension embraces two tasks providing two samples of the target language to be read. The tasks are not alike in terms of their topic and text type. The first task covering 10 items (assessment units) is usually an objective one with discrete scoring, in most cases it is a banked gap-filling or a multiple-choice test. The second task is a set of open-ended questions (10 items) that checks the candidate's comprehension of the text. Evidently these WH-questions (as a form of question) are familiar for learners as natural and hands-on tools that have been applied in language teaching for centuries (Kelly 1969: 135). However, the concrete questions' references in texts of level B2 cannot be easily associated with the bits of information or easily pinpointed within the text. The questions might address not explicit, not clear-cut but somewhat vague and hidden piece of information for the L2 reader, that is, answers that are implied by the utterances.

Nevertheless, the set of the 10 questions in the task may appear as guidance for the reader to explore the informational structure of the text because their succession follows the course of the text. Therefore, when answering the 10 open-ended questions the candidates' task is twofold: firstly to find the location of the piece of information (scanning for information) that is asked about by the question and secondly on the basis of text location to answer the question.

Query

The chosen text was about the brief history of genealogical investigations of Hungarian people in the 19th and 20th century. The text explores why these investigations to trace family trees were always popular: they used to be means of proving one's noble ancestors to gain privileges. The 400-word long text contains vocabulary of historical character (e. g. *ősök, jogok, kiváltságok, polgári származás, levéltár, nemesség-adományozás, kisnemes* etc.) meeting the requirements of fairly elaborate texts aligned to level B2 (Szabó 2011).

Having come across with the exam task based on this text the candidates were challenged by questions whose lexical formation was not in the least identical with the words by which the informational content was assembled in the text. That is no literal overlap but only lexical parallelism might be observed between the text and the task. This lexical difference refrains the candidate from lifting certain literal content fragments into his/her answer on the basis of superficial lexical similarity but without comprehending the text (which is the major goal of the reading test).

In my query I have analyzed 11 candidates' answers given to the 10 questions and assessed them following the key provided by the test constructors. The goal of the analysis was to explore the thoroughness of text comprehension of candidates and learn what role coreferential links might play in this process.

Findings

If the candidate's answer (its content) meets the expected answer elaborated by the key, it is assessed as a successful item response, otherwise it is incorrect. The results were as follows:

Item	Scores (correct: grey; incorrect: white)										
1.	grey	grey	grey	white	grey	white	grey	grey	grey	white	grey
2.	grey	grey	grey	grey	grey	grey	grey	grey	grey	grey	grey
3.	white	grey	grey	grey	white	white	grey	grey	grey	grey	white
4.	white	grey	white	white	grey	white	white	white	white	white	grey
5.	white	white	grey	grey	white	grey	white	white	grey	white	white
6.	white	white	grey	white	grey	white	white	white	white	grey	white
7.	grey	white	grey	white	grey	white	grey	white	grey	grey	white
8.	white	white	grey	white	grey	white	white	white	white	white	grey
9.	white	grey	white	white	grey	grey	white	white	white	white	grey
10.	white	grey	grey	white	grey	grey	white	white	white	grey	grey

As can be seen from the chart the first two questions (1. Mi állhatott régen a genealógiai bizonyítékok megszerzésének hátterében? (min 2.) 2. Milyen műhelyek jöttek létre a kezdeti nagy érdeklődés miatt?) are not easier to be answered than the succeeding questions. Yet the

238

majority of the candidates managed to provide a plausibly correct response. The relatively high number of the correct answers (especially in the case of the second answer) to the first two questions might account for the initial position of the answers' reference within the text: presumably the candidates could find the links between the first two questions and the first text paragraph better than further in the text. The beginning of a text in most cases is more salient and transparent for the reader because its content does not refer back to former text locations to be comprehended by the reader.

Question no. 3 (*3. Minek számított a családfakutatás a második világháború után?)* involves a verb–complement structure that needs sophisticated grammatical knowledge of the candidate. However, the time adverbial helps the reader highlight the location of the piece of information needed to the correct answer.

The next question proved to be the most difficult (*4. Milyen többlettudás jellemzi a valódi családfakutatót?*) and this fact might have contributed some learners' losing their way in the text (as it can be seen in the visual arrangement of the inappropriate answers: only three out of the eleven candidates gave a correct reply).

This quantitative matrix of the correct and wrong answers can provide only a limited insight into the candidates' item-response strategies, specifically how coreference may help in providing correct answers.

Coreference and coreference of open-ended questions

According to Halliday (Halliday 1976: 3) if "two items are identical in reference", they are coreferential (common reference). In a text to establish common reference certain words are repeated, sometimes synonyms are used instead or in the case of nouns, adjectives, especially superordinate or general equivalents might be used. Verbal, clausal or pronominal substitutions also might appear to carry the same reference as former text parts. Most frequently anaphoric and

cataphoric coreference are at hand at first place for the speaker to refer to something appearing further in a text or referring back to a former element respectively to build up cohesive ties between text parts.

In open-ended questions the question word and its complements establish a coreferential link to the content expected in the response. As a matter of fact it is the question word itself that does elaborate the semantic content of the verb (Langacker 2008: 198) missing from the utterance, though it elaborates it only to a limited extent, only by the schematic reference of the interrogative pronouns. Yet it is the part of the speech that despite its schematic character finally narrows the focus of the question. The exact reference of the interrogative pronoun must be tracked by the candidate who answers the question despite the opacity of the question word. The more (s)he comprehends the interrelated content pieces, the less opaque is the question.

To take a qualitative view of the first four (both correct and wrong) answers given by the candidates it can be seen that many of them could not reply to the question word at all:

Question 1:
Mi állhatott régen a genealógiai bizonyítékok megszerzésének hátterében? (min 2.)
Answers:
√ csak a jogok és kiváltságokat megszerezni kell
√ a jogok és kiváltságok
√ jogok és kiváltságok megtartása
* akadtak szép számmal
* az emberek kíváncsik voltak a család történetéről

Question 2:
Milyen műhelyek jöttek létre a kezdeti nagy érdeklődés miatt?
Answers:
√ megalakultak kutatóirodák
√ családtörténeti kutatóirodák
√ családtörténeti kutatóiroda
√ családfakutató
* szászedik
* kalandos emberekre, akik kutatnak

Question 3:
Minek számított a családfakutatás a második világháború után?
Answers:
√ úri passziónak számított
√ úri passziónak
√ úri passziónak bélyegzett
* családfakutatás az egy amatőr kutató
* az érdekségnek
* bélyegek
* hogy alakutja a családtörténet kutatóivá
* nemesi származását bizonyítasnak
* a viták hogy ki hová tartozott

Question 4:
Milyen többlettudás jellemzi a valódi családfakutatót?
Answers:
√ íráskép elolvasásai tudásak
√ történelmi és kalligráfiai
√ írásképek elolvasás, tudni latinul németül

The answer to the first open-ended question was triggered by the *Mi?* [what?] question word thus nouns (actually 2) were expected: *jogok és kiváltságok*. One of the wrong answers reveals that not even the expected part of the speech was recognized by the learner, at least according to his/her answer. The other wrong answer was given not on the basis of the text and in this respect it was irrelevant though the question word's nominal reference was correctly recognized. In the case of the fourth question that proved to be the most difficult one (together with the eight one according to the chart of scores), only three candidates could give a proper answer and the rest of the candidates could not give any kind of reply. In the case of the fourth question supposedly not only the coreferential content was missed being identified by the candidates. Presumably neither the location of the bit of information expected as an answer to the question was found, nor the question itself might have been well understood: *Milyen többlettudás jellemzi a valódi családfakutatót?* The comprehension of this question adheres to the comprehension of the whole text. This latter question conveys several presuppositions: 1. There were genealogists in Hungary. 2. There were 'real' and 'not real'

genealogists and 3. the true (eager) ones 'did have' some extra knowledge to trace up family trees. Only these pieces of information (in the form of presuppositions in the question) together with the former context could have helped the candidates to an answer this question.

Conclusion

We may not be right when guessing language learners' strategies applied answering open-ended question. But as test constructors' in a quantitative and qualitative approach we have to make an attempt to decipher their potential problems. As seen formerly from a very limited number of questions and some wrong answers given by the candidates coreference plays an essential role in reading comprehension. For learners of Hungarian as L2 besides the macro elements of coreference such as synonymic parts of speech or chunks of the texts, micro elements (inflectional morphemes, pronouns with schematic reference) should also be taken into consideration. On higher levels of linguistic proficiency these tools are crucial to learners' cognitive effort to be able to paraphrase, reformulate, imply, conclude information pieces of the text that are not salient for them for the first sight.

References

ECL nyelvvizsga. http://ecl.hu/ Zugriff: 31.03.2015.

Halliday, M.A.K, Hasan, R.: *Cohesion in English*. London 1976.

Hegedűs, Rita 2003: Feladat, gyakorlat – nyelvvizsgáztatási tapasztalatok a magyar mint idegen nyelv területén. In: *Hungarológiai Évkönyv.* IV. évf. 1. sz. (2003), 32–43.

Huszti, Judit: Kommunikatív szempontok a nyelvvizsgán. In: Kárpáti, Eszter/Szűcs, Tibor (eds.): *Nyelpedagógia.* (2002) 166–173.

Kelly, Louis, G.: *25 centuries of language teaching: an inquiry into the science, art, and development of language teaching methodology, 500 B.C.–1969.* Rowley, Massachusetts 1969.

243