

EWA WITALISZ, JUSTYNA LEŚNIEWSKA
Instytut Filologii Angielskiej

PERCEPTION OF WORD DIFFICULTY IN L2: TEACHER VS. LEARNER JUDGMENTS

1. Introduction

Word difficulty in L2 is typically discussed in the context of vocabulary acquisition, namely, what makes words difficult or easy to learn. Considering the common educational settings, specifically the context of the assessment of learner written output, it is interesting to study which individual lexical items are perceived to be difficult and to what extent such perception is similar among learners and teachers.

In qualitative writing assessment, vocabulary assessment plays a significant role and is quite a challenging task for raters. Whether they use analytic or holistic assessment scales, raters are typically expected to evaluate the vocabulary in a learner text in terms of range (size), accuracy (errors) and lexical sophistication. Of these three, it is lexical sophistication that causes particular difficulties. After adequate training, raters can agree on what constitutes a wide range of vocabulary in a given set of learner texts and in which text the vocabulary can be seen as repetitive. Lexical accuracy, once a particular language norm is accepted, should be even less problematic. However, assessing vocabulary in terms of which words should be classified as advanced or sophisticated depends on numerous judgments, which are likely to be affected by one's own individual perception of word difficulty.

This paper briefly presents the study of the perception of word difficulty conducted in Kraków and Tarnów in 2007.

2. Subjects and procedure

In order to compare teacher and learner judgments of word difficulty, we asked teachers and learners to assess the vocabulary in a sample of learner written English. There were three groups of subjects. Group 1 consisted of 14 teachers of English (Jagiellonian University, English Department, Kraków) with at least 5 years' experience as EFL teachers and examiners. Group 2 comprised 57

graduate students of English (Jagiellonian University, English Department, Kraków), while Group 3 consisted of 62 undergraduate students of English (PWSZ Tarnów). Although we considered both groups of students as learners in our study, they differed with respect to their level of English, the Kraków group being more advanced,¹ and their involvement in teacher training and EFL teaching itself. As opposed to the Tarnów students, all the Kraków students had received extensive teacher training and had had some experience of teaching, either as trainee teachers during their teaching practice or already having jobs as teachers of English. However, they were still studying English for their final exam, where their English was to be assessed, which means that they still had the learners' perspective on vocabulary acquisition.

The subjects were asked to assess the vocabulary in the following text produced by an advanced EFL learner:

The situation of a large part of Polish society is becoming more dramatic every day. Because of the period of economic changes, unemployment has risen rapidly and the salaries have gone down, which has caused tremendous poverty. Many people are not able to improve their situation no matter how good intentions they have and what choices they make. For inhabitants of large cities things don't look so pathetic. People from smaller communities are the ones whose situation is the worst, especially regarding youth or families with many offspring. These children are often simply hungry, not to say anything of their living conditions. Their lifestyle is a far cry from life of the young people from rich families, being raised in wealth and luxury. That is why a vast number of poor children are ashamed of their status and encounter signs of contempt from the other social groups. The prejudice against the poor is rising and they are considered as the worse part of society. The social organizations which are supposed to help such people, in fact don't have sufficient means for that and the social benefits are becoming lower, while the need for them is still rising. The poor aren't able to improve their situation on their own, because their walk of life doesn't enable them to do it. It is like a vicious circle: the children of poor parents don't have a chance of breaking through the social barriers and in future their lifestyle will be the same as it is now.

In Task 1, the learners were to highlight the words which they considered difficult, while the teachers were to imagine that they were assessing the vocabulary in that text and they were asked to highlight the words which they would appreciate as advanced or sophisticated. In Task 2 the subjects were to explain why they found particular words difficult (learners) or advanced/sophisticated (teachers). The instructions for the learners were also given in Polish (L1), and they were encouraged to provide answers in Polish so that potential problems with L2 would not affect the content of their responses.

¹ Tarnów students' level of English – B2 (FCE).

Kraków students' level of English – C1-C2 (CAE-CPE).

For the teachers, who were also trained examiners, the task was simply a single component of qualitative writing assessment. Instead of assessing the learner text globally, taking into consideration such criteria as task fulfilment, cohesion and coherence, range and accuracy of grammatical structures, etc., they were to focus selectively on one aspect of vocabulary, namely lexical sophistication. In the case of the learners, the data collection method was a metalinguistic knowledge task, in which "instead of direct elicitations of language, learners are presented with linguistic stimuli, in reaction to which they must make other active decisions, ratings, comparisons, and revisions about the form or meaning of the stimuli" (Chaudron 2003: 796). It is important to note here that although metalinguistic knowledge tasks, e.g. grammaticality judgment tasks, are widely used and their validity has been confirmed by numerous studies, their findings have also been found inconsistent with other measures of subjects' productive capacities (Chaudron 2003: 805–806). In our study this may mean a discrepancy between the subjects' perception of difficulty and their actual knowledge of the word, e.g. the subject may rate the word as easy although he/she cannot really use it well or does not know it at all. As will be shown in the discussion of the results, the weakness of the method may substantially affect the results.

3. Research questions and hypotheses

Our basic research question was whether the teacher and learner judgments of word difficulty were similar in terms of the words identified as difficult or sophisticated and in terms of the criteria used. We were also interested to what extent each group of the subjects would be unanimous in their judgments. On the one hand, it was reasonable to expect similarity between the groups and unanimity within the groups as a result of all subjects being part of the same teaching process and education system, with the teachers being more specific in the justification of their judgments. On the other hand, it was also reasonable to expect that the perception of word difficulty may vary between the subjects because it is strongly affected by individual learner differences² and one's teaching experience. Even trained examiners may have different levels of sensitivity to error, and in the case of academic teachers there might also be some bias resulting from their academic interests. For instance, literature teachers may appreciate some literary devices in a learner text, while linguists may appreciate words for their etymology.

4. Criteria for word difficulty

To analyze the qualitative data elicited in Task 2, where the subjects were to justify their judgments of difficulty in their own words, we classified their responses into the following categories of criteria for word difficulty:

1. intralexical

² For example, if pronunciation is particularly difficult for a learner, then this is likely to be his/her major criterion for word difficulty.

2. intralingual (word frequency)
3. interlingual (cross-linguistic influence)
4. syllabus constraints.

Our selection of intralexical criteria was based on Laufer's (1997) well-known paper "What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words." This is how Laufer's criteria for word difficulty may be illustrated with the words from our text:

1. length: *tremendous*
2. spelling: *tremendous*
3. pronunciation: *prejudice*
4. sound-script incongruence: *prejudice*
5. morphological complexity
 - inflectional: *risen, offspring*
 - derivational: *unemployment*
6. confusability
 - synforms: *economic/economical, enable/unable, raise/rise*
 - deceptive morphological transparency: *offspring* (off/spring), *encounter* (en/count/er)
7. semantic features
 - specificity: *sufficient means* (for enough money)
 - idiomaticity: *vicious circle, far cry, walk of life, break through*
 - register restriction: *offspring, encounter* (formal)
 - multiple meaning: *far cry* (a computer game), *offspring* (a popular American punk rock band).

Word difficulty may also result from the relationship of a particular word to other words in the same language (intralingual criteria), which shows in word frequencies. To identify the rare words in our text we used *Collins COBUILD* frequency bands from *Collins COBUILD Dictionary for Advanced Learners* (2001) presented in Table 1.

Table 1. Collins COBUILD (2001) frequency bands

◆◆◆◆ band 1	ca. 680 most frequent words	75% of the language	95 % of the language
◆◆◆◆ band 2	next 1040 most frequent words		
◆◆◆ band 3	next 1580 words		
◆◆ band 4	next 3200 words		
◆ band 5	next 8100 words.		
remaining			5%

The least frequent words that appeared in our text turned out to be *offspring*, *worst*, *worse*, *vicious circle* (band 5) and *contempt*, *prejudice*, *pathetic* (band 4). As has already been mentioned, the word *offspring* may be difficult because of

its other meaning, possibly more familiar to the learners, and this is also confirmed by its frequency: while in its primary meaning ('children') it belongs to low frequency words in the *Collins COBUILD* corpus, it has a very high Internet frequency (over 22,000,000 Google hits) resulting from the other meaning ('American punk rock band').

Other reasons for word difficulty spring from the relationship between a particular word and the words in other languages (interlingual criteria). An example from our text is the word *pathetic*, a false friend of the Polish word *patetyczny*.

Finally, words can be perceived as difficult due to their place in the syllabus: they are seen as more difficult synonyms of the words introduced earlier, e.g. *offspring* (*children*), *tremendous* (*great, big*), *rapidly* (*quick, fast*), *sufficient* (*enough*).

To conclude, of all the words that appear in our text the word *offspring* could be expected to be identified as difficult or sophisticated for a number of reasons. It is morphologically complex (the same singular and plural form); it is morphologically deceptive (learners' knowledge of the morphemes *off* and *spring* interferes with its meaning); it has multiple meanings and its other meaning may be much stronger in learners' mental lexicons; it is a low frequency word, and it is introduced late in the syllabus. Interestingly, the words *worse*, *worst* also meet some criteria for word difficulty. They are morphologically complex (irregular forms) and they have equally low frequency (*Collins COBUILD* band 5), but they are typically introduced very early in the syllabus, so whether they are perceived as difficult or sophisticated depends on which criteria are given priority.

5. Results

The most unexpected result in our study was the response of the Tarnów group. While the average number of words identified as difficult or sophisticated was ca. 11 in the group of Kraków students and in the group of teachers, the Tarnów average was only 3.5 words per text, with as many as 11 students out of 62 claiming that there were no difficult words in that text. Considering their lower level of English, this answer was quite surprising. After all, one could assume that a less proficient student would identify more, rather than fewer, difficult words. This brings back the question of the validity of the metalinguistic knowledge task. As a method of eliciting data from learners, it simply failed in the case of some Tarnów students. This may have resulted from their high anxiety, fear of negative evaluation, an attempt to boost their self-esteem by pretending that all those words were easy, or simply from their general inability to comment on their own experience of second language acquisition. Although all tests were administered in the same way,³ some Tarnów subjects must have mistaken the task for another language test, where their linguistic performance would be assessed, and our request to share their learning experience was mistaken for a vocabulary test question.

³ The tests were anonymous and the subjects were assured that they were taking part in a study. Their questions were answered and explanations were provided whenever any uncertainty appeared.

The results were grouped in two categories: the words selected by the most subjects in each group (Table 2) and the prevailing criteria used in each group (Table 3).

Table 2. Words ranked by the percentage of subjects who identified them as difficult/sophisticated

	Teachers			Kraków students			Tarnów students	
1	offspring	93%	1	tremendous	96%	1	tremendous	47%
2	tremendous	79%	2	offspring	86%	2	pathetic	42%
3	(a) far cry (from)	79%	3	pathetic	77%	3	encounter	31%
4	(a) vicious circle	79%	4	prejudice (against)	68%	4	contempt	27%
5	walk of life	64%	5	(a) vicious circle	65%	5	prejudice (against)	26%
6	pathetic	57%	6	a far cry (from)	63%	6	(a) vicious circle	21%
7	encounter	57%	7	encounter	61%	7	vicious	16%
8	prejudice (against)	50%	8	inhabitants	53%	8	offspring	13%
9	contempt	43%	9	contempt	46%	9	risen	11%
10	breaking through	43%	10	sufficient	35%	10	a far cry (from)	10%

As to the selection of words, out of the top 10 words as many as 7 appear in all groups: *vicious circle*, *contempt*, *encounter*, *offspring*, *pathetic*, *prejudice*, *tremendous*, indicating a strong similarity between the groups. It is also interesting that the word *tremendous* is at the top (first or second) in all three groups.

Table 3. Criteria for word difficulty ranked by the number of indications in each group

	Teachers	Kraków students	Tarnów students
1	syllabus constraints	frequency	frequency
2	idiomaticity	idiomaticity	pronunciation
3	register restrictions	syllabus constraints	confusability
4	collocability	register restrictions	UNFAMILIAR
5	Latin origin	spelling	idiomaticity
6	specificity	false friend	difficult to remember
7	morphological complexity	length	spelling
8	frequency	confusability	false friend
9	phrasal verb	morphological complexity	no associations
10	literary	collocability	length

The list of the prevailing criteria used in each group shows how the Kraków group of graduate students with some teaching experience understandably shares some similarity with the Tarnów group of learners and some with the group of teachers. The top criterion for both groups of learners was frequency: words are difficult because they are rare. One may wonder, though, if the learners really meant word frequency in the language as shown by corpora studies, or possibly

word frequency in the teaching materials they were familiar with, which brings this criterion close to syllabus constraints. Not surprisingly, the teachers' top criterion was pedagogical and closely linked to the educational setting. Words are difficult because learners acquire them later, so using a particular word indicates the level of advancement that the learner has reached. It is interesting how this perspective is quite contrary to that of syllabus designers, for whom learnability is one of the criteria for sequencing the lexical component of the syllabus, i.e. easier words should be taught earlier (White 1988: 48–50), which means that there are some tangible criteria for word difficulty determining their place in the syllabus, rather than the other way round.

While word frequency is the top criterion that the Kraków students share with the Tarnów group, their next three criteria are exactly the top criteria in the teachers' group: idiomaticity, syllabus constraints and register restrictions. The Tarnów group, on the other hand, is much more distinct from the other two groups with respect to the criteria than the selection of words. Their second criterion for word difficulty was pronunciation, which was again quite an unexpected result because the subjects were dealing with a sample of written English. The Tarnów students' concern about pronunciation may have resulted from the emphasis in the instruction they had received as well as their own personal difficulties in that area of the language. What was again unexpected was that some words were perceived by the Tarnów group as difficult because unfamiliar. Naturally, we wanted the subjects to pass judgment on the words that were familiar to them and simply did not expect that advanced learners would have problems with the vocabulary in that text.

Finally, as regards the question of unanimity in the group, or in other words, inter-rater agreement, the Tarnów group is again distinct from the other subjects. While the Kraków students and teachers demonstrated a high level of inter-rater agreement in the selection of the words – the top three words were accepted by over 75% of subjects in the group, the Tarnów students did not even reach 50% in the number of subjects accepting the top word, which means that there is not a single word that would be perceived as difficult by the majority of the group. Their lack of agreement also shows in the discrepancy between refusing to identify any difficult words in the text on the one hand, and marking words that were unfamiliar on the other. Not surprisingly, neither of these was done by a single subject from the other two groups.

6. Conclusions

As regards the criteria for word difficulty, the results of our study are quite consistent with Laufer's (1997) results concerning intralexical factors affecting word learnability. The difficulty-inducing factors she identified were all indicated as criteria for word difficulty by the subjects in our study:

- sound-script incongruence
- inflectional complexity
- derivational complexity
- deceptive morphological transparency

- synformy
- specificity
- register restrictions
- idiomaticity
- multiple meaning.

The factors which she found to have no clear effect were either mentioned by a minority of subjects (word length, abstractness) or never mentioned (part of speech).

Our research questions, however, basically concerned the comparison of teacher and learner judgments of word difficulty, and the following conclusions may be drawn for the analysis of the results:

1. Despite some differences in the criteria used, there is a strong similarity in the lexical items identified as difficult or sophisticated between all three groups.
2. The Kraków group of graduate students could be seen as a transition group between learners and teachers. They display some similarity to the Tarnów group (the top word selected – *tremendous* and the top criterion – word frequency) and a strong similarity to the teachers' group in terms of the other top criteria and high inter-rater agreement.
3. The Tarnów group, although apparently strongly integrated, provided diverse data, which without further investigation of their individual differences cannot be easily interpreted. It is possible, though, that in the case of some Tarnów subjects the data elicitation method was inadequate and a different method, less reminiscent of a testing situation, would have proved more valid.
4. As to the differences between the teachers, there were some criteria resulting from their individual academic interests: some words were considered sophisticated by a teacher of linguistics because of their Latin origin; a phrase (*risen rapidly*) was appreciated for alliteration, an expression (*breaking through the social barriers*) was appreciated for its metaphorical quality by literature teachers. However, these differences were minor in comparison to the high inter-rater agreement concerning the top words and criteria.

References

- Chaudron C. (2003): *Data Collection in SLA* [w:] C.J. Doughty i M.H. Long red. *The Handbook of Second Language Acquisition*, Oxford, s. 763–828.
- Collins COBUILD English Dictionary for Advanced Learners, London 2001.
- Laufer B. (1997): *What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words* [w:] N. Schmitt i M. McCarthy red. *Vocabulary: Description, Acquisition and Pedagogy*, Cambridge, s. 140–155.
- White R. (1988): *The ELT Curriculum*, Oxford.

Streszczenie

Odczuwanie stopnia trudności słownictwa w języku obcym. Poglądy nauczycieli a poglądy uczących się

Artykuł przedstawia wyniki badania odczuwania stopnia trudności słownictwa w języku obcym, przeprowadzonego wśród nauczycieli języka angielskiego i uczących się na poziomie zaawansowanym. Celem badania było ustalenie, w jakim stopniu nauczyciele i uczący się postrzegają te same słowa jako trudne lub świadczące o zaawansowaniu językowym oraz jakie stosują kryteria. Badani otrzymali tekst napisany przez osobę uczącą się języka angielskiego na poziomie zaawansowanym i ich zadaniem było zaznaczenie słów trudnych oraz podanie kryteriów wyboru. Kryteria trudności słownictwa zostały podzielone na intraleksykalne (długość słowa, pisownia, wymowa, złożoność morfologiczna, mylące podobieństwo form, cechy semantyczne, jak wieloznaczność, ograniczenia rejestru, idiomatyczność), intralingwalne (częstotliwość słowa w języku), interlingwalne (interferencja innego języka) oraz dydaktyczne (typowa sekwencja słownictwa w programach nauczania). Badanie wykazało, iż pomimo różnic w zastosowanych kryteriach trudności słów, wybór słów uznawanych za trudne przez uczących się w dużej mierze się pokrywał z wyborem słów uznanych przez nauczycieli za skomplikowane.