

Designing a Self-Assessment Instrument for Developing the Speaking Skill at the Advanced Level

Aleksandra Jankowska and Urszula Zielińska

Abstract The present paper describes an attempt at developing in college students the ability to self-evaluate their own recordings of short speeches on topics of their choice, made as a credit requirement for speaking classes. The students were asked to record short speeches (3–5 min) and hand them in. Initially, no self-assessment was required and then students were asked to evaluate their recordings, but no guidelines were provided. Despite the fact that all the students took a semester course in learner-training and that techniques for implementing elements of self-assessment in their own teaching were discussed during methodology classes, very few students manifested the ability (or willingness) to evaluate their speeches, limiting the evaluation to general remarks such as “I made some grammar mistakes” or “I need to work on my pronunciation”. As a result of this experience, a short self-assessment instrument based on the *Common European framework of reference* (CEFR) descriptors for the speaking ability (spoken production) at the C1 level was developed and the students were asked to use it to evaluate their speeches. Selected students were then interviewed in order to find out their opinions on the usefulness of the instrument in helping them to identify their strengths and weaknesses in speaking. Both the students’ and the teachers’ opinions were later analyzed and the results of this analysis will serve as a basis for developing an improved version of the self-assessment sheet.

1 Introduction

Contemporary approaches to language teaching, such as the communicative approach or task-based teaching, see the development of effective oral communication as one of their main goals. There is no agreement, however, as to how this

A. Jankowska (✉) · U. Zielińska
Adam Mickiewicz University, Poznań, Poland
e-mail: ajank@amu.edu.pl

U. Zielińska
e-mail: urszula.zielinska12@gmail.com

goal is to be achieved and we are far from fully understanding the processes underlying speaking. According to Levelt (1989), speech production consists of four main stages: *conceptualization*, *formulation*, *articulation* and *self-monitoring*. During the conceptualization stage the content of the message is planned. This content is then matched to appropriate words and phrases which are connected following the rules of grammar and encoded phonologically. This pre-planned utterance is then physically produced during the stage of articulation. The whole process is controlled by a monitor, which is a part of the conceptualizer and is active both during and after articulation. For the process of speech production to be fluent, a degree of automaticity is necessary. The analysis of speech in terms of fluency, accuracy and complexity (Skehan 1998) shows that learners find it difficult to concentrate on all of these factors simultaneously during speech production and, depending on the nature of the task, focus on one of them at the expense of the other two. Additionally, the quality of students' output has been found to be positively influenced by such factors as planning time and opportunities for task rehearsal and repetition (cf. Bygate 1996, 1999). In view of the above, it seems reasonable to suggest that planning, recording and then analyzing one's speech may be beneficial to foreign language learners wishing to improve their oral skills.

2 Testing and Assessing Speaking

In the literature on language learning and teaching the terms *testing* and *assessment* are used when talking about measuring learners' progress and proficiency. Testing usually refers to the more formal ways of checking students' knowledge and is a subset of assessment, which is an ongoing process of providing students with feedback on their performance. Most of the techniques used for assessment can also be used in testing and that is why in the remainder of the paper these two terms will be used interchangeably. Another important distinction is that between *formative* and *summative assessment*, where formative assessment refers to "evaluating students in the process of "forming" their competencies and skills with the goal of helping them to continue the growth process" (Brown 2004, p. 6), while summative assessment "aims to measure, or summarize what a student has grasped" (2004, p. 6).

Assessing and testing speaking is difficult and time-consuming. The difficulty is mainly due to the fact that speech is temporary and a teacher/assessor needs to conduct assessment immediately at the time the student is talking and very often rely on his/her memory to provide an accurate evaluation and feedback. Nowadays, the use of technology can help solve this problem with recordings of students' oral performance constituting a viable, although still rarely used, option. Assessment of speaking skills is also an extremely subjective process with many factors influencing the teacher's judgment. This problem can be minimized by developing and following clear scales. Many such scales are already available and their strengths and weaknesses are discussed in the literature on the subject (e.g. Luoma 2004; Hughes 2011). The major problems connected with assessing speaking are best

summarized by Alderson and Bachman in their preface to Luoma (2004, pp. iv–v) volume on the subject:

Speaking is (...) the most difficult language skill to assess reliably. A person's speaking ability is usually judged during a face-to-face interaction in real time, between an interlocutor and a candidate. The assessor has to make instantaneous judgment about a range of aspects of what is being said, as it is being said. This means that assessment might depend not only upon which particular features of speech (e.g. pronunciation, accuracy, fluency) the interlocutor pays attention to at any point in time, but upon a host of other factors such as the language level, gender, and status of the interlocutor, his or her familiarity to the candidate and the personal characteristics of the interlocutor and candidate.

Speaking is a complex skill and its assessment includes the assessment of such areas as pronunciation, grammar, vocabulary, fluency, comprehensibility, coherence and cohesion, as well as the ability to interact and adjust one's speech to a particular social context. These areas are reflected in most rating scales used to assess oral proficiency. Brown's (2001, pp. 406–407) oral proficiency scoring categories include grammar, vocabulary, comprehension, fluency, pronunciation and task, TOEFL speaking rubrics are divided into delivery (fluency, intonation, rhythm, pronunciation), language use, divided into vocabulary (which is evaluated with reference to diversity, sophistication and precision) and grammar, evaluated on the basis of range, complexity and accuracy, and topic development, assessed taking into account coherence, idea progression and content relevance (Hughes 2011, p. 99), while the criteria for the IELTS speaking test include fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation (Hughes 2011, p. 104), to give just a few examples.

In designing assessment tasks, we should also take into account the various types of talk we engage in. Brown and Yule (1983) distinguish four different types of information talk: *description*, *instruction*, *storytelling* and *expressing and justifying opinions*. Bygate (1987) differentiates between *factually-oriented talk* including description, narration, instruction and comparison, and *evaluative talk* comprising explanation, justification, prediction and decision. It is important that all the above types of talk are included in assessment procedures.

Tasks used for the purposes of assessing speaking can be grouped into several categories, one of which was put forward by Brown (2004, pp. 141–142), who lists the following types of speaking performance that can be the focus of assessment:

- (1) *imitative*, in which students are asked to repeat short words or phrases and whose aim is to focus on pronunciation;
- (2) *intensive*, which include reading aloud or sentence or dialogue completion;
- (3) *responsive*, which take the form of very short interactions;
- (4) *interactive*, which are extended versions of responsive tasks;
- (5) *extensive*, which “include speeches, oral presentations, and story-telling, during which the opportunity for oral interaction from listeners is either highly limited (perhaps to non-verbal responses) or ruled out altogether. Language style is frequently more deliberate (planning is involved) and formal for extensive tasks” (Brown 2004, p. 142).

Specific examples of tasks used for the purpose of assessment include, among many others, imitation, interviews, picture descriptions, role plays and simulations, collaborative tasks, discussions, and live and recorded monologues (Thornbury 2005, pp. 125–126; Johnson 2008, p. 319).

Students can be asked to perform the tasks individually with the teacher/assessor acting as an interlocutor, in pairs or in groups, depending on the type of task and the aim of the test. Individual testing is time-consuming and stressful due to the unequal balance of power between the tester and the examinee, but it allows for flexibility in approaching each candidate. Another weakness of this type of arrangement is a limited number of types of tasks which can be employed. Both pair and group work allow for more variety in this respect, although they are also not without weaknesses, the major one being the influence of each candidate's proficiency level and personality on the performance of the other members of the group (Luoma 2004, pp. 35–41).

Taking into account the fact that most speaking is interactive in nature, it is not surprising that the most common assessment/testing techniques try to emulate that feature. There is still, however, room for monologic tasks during which students are given an opportunity to practice longer stretches of discourse (Luoma 2004, p. 44; Thornbury 2005, p. 126). Monologic tasks are used in the speaking part of the “iBT/New generation TOEFL” test as well as in the IELTS speaking test (Hughes 2011, pp. 99–103).

Typically, assessment is conducted by the teacher, but this is not the only option and both peer- and self-assessment should be considered. Peer-assessment is closely related to principles of cooperative learning and “is simply one arm of a plethora of tasks and procedures within the domain of learner-centered and collaborative education” (Brown 2004, p. 270). Self-assessment will be discussed in more detail in the following section of the paper.

3 Self-Assessment

Self-assessment might be considered by some researchers and practitioners as an “absurd reversal of politically correct power relationships” (Brown 2004, p. 270). However, to those who adhere to less conventional ways of teaching, this notion is extremely valuable, because it is so closely connected with the concept of developing autonomy and self-regulation. It would be very difficult to imagine independent, successful learners without the skill and the willingness to reflect on their performance and introduce adjustments into their own ways of learning a foreign language. If self-regulation is expected to be developed and improved, then the three subprocesses, namely forethought, performance or volitional control, and, most importantly, self-reflection (Zimmerman 2000) would be incomplete without self-judgment and self-evaluation. Once observed, analyzed and evaluated, different aspects of one's own performance, whether oral or written, become the foundation for a change, through which specific goals of an individual are attained. In other

words, by looking back and assessing performance, a person judges the effectiveness of techniques employed in learning the language and can thus adjust and modify their actions. Furthermore, “our regulatory skill, or lack thereof, is the source of our perception of personal agency that lies at the core of our sense of self” (Zimmerman 2000, p. 13). It would be impossible, then, to achieve this state of personal agency without the ability to self-regulate, self-reflect and self-assess.

Brown lists self- and peer-assessment among the “best possible formative types of assessment and possibly the most rewarding”. The five categories he distinguishes include (2004, p. 270):

- (1) assessment of (a specific) performance;
- (2) indirect assessment of (general) competence;
- (3) metacognitive assessment (for setting goals);
- (4) socioaffective assessment;
- (5) student-generated tests.

The first type of assessment requires an immediate (or at least not delayed) evaluation of the performance and is usually based on a checklist or some other defined scale. Journals and video-recordings are also used for that purpose. It was this type of self-assessment that had become the focus of our interest.

4 The Study

The present paper reports the results of an action research project in progress. Action research is operationalized here as “a form-of self reflective enquiry undertaken by participants in social situations in order to improve the rationality and justice of their own practices, their understanding of these practices, and the situations in which these practices are carried out” (Carr and Kemmis 1986, pp. 220–221, as quoted in Nunan 1989, p. 12).

In designing the study, the authors followed the procedure put forward by Kemmis and McTaggart (1989), cited in Nunan (1989, p. 12) in which the following stages of action research are identified:

- *Phase I*: Develop a plan of action to improve what has already been happening.
- *Phase II*: Act to implement the plan.
- *Phase III*: Observe the effects of action in the context in which it occurs.
- *Phase IV*: Reflect on these effects.

The aim of this research project was to develop a self-assessment checklist to help students evaluate their speaking skills, to evaluate it and, if necessary, suggest changes in the design of the instrument to be implemented during a follow-up stage.

4.1 Participants

Forty-six students who were involved in the initial stages of the study were 3rd year students of English of a teacher training college. Their course in English as a foreign language included classes in grammar, writing, speaking integrated with reading, and speaking integrated with listening. During the first two years of study they also had separate classes devoted to pronunciation practice. As part of the teacher training component of their study program (year 1 and 2), the students had classes in methodology devoted to ways of teaching all the aspects and skills of the target language as well as assessment and testing. Additionally, during the first year, they underwent a semester course in learning strategies aiming at improving their own ways of working on language development. The course also included elements of self-assessment.

The students were accustomed to recording their oral performances as they were asked to do it for their first year listening/speaking course. The speeches were to be recorded once a month and be 3–5 minutes long. The students were allowed to choose their own topics but they were encouraged to talk about the issues discussed during classes. In the second year, the students were not asked to submit recorded speeches and their speaking ability was assessed on the basis of their in-class performance including presentations, and a mock exam conducted at the end of the year.

4.2 Stage One: Preliminary Assessment

In the first semester of the academic year 2010/2011, the students were required to record 3 speeches *per* semester, each 3–5 minutes long. The students were able to select topics they wanted to address but were encouraged to talk about topics discussed in class. The students recorded their speeches using a variety of devices and software, and then submitted them on a CD or by e-mail. One student recorded her speeches on an audio cassette. The teacher then listened to the recordings, made notes on them and provided oral feedback on the students' performances during classes. The students were asked to present the main points of their speeches to their classmates. Once during the semester the teacher held individual conferences with the students in order to discuss their speeches in private in more detail. Providing feedback turned out to be quite difficult mainly because of the time required to do so effectively. This and the fact that the students were going to end their formal education soon constituted an incentive for the teacher to introduce elements of self-assessment into the project.

4.3 Stage Two: Preliminary Self-assessment

During the second semester, the students were asked to record the same number of speeches of the same length; this time, however, they were to attach a transcript of their speech together with phonetic transcription of its fragment and a short written

evaluation of their performance. In order to help them with the task, the teacher conducted a discussion session in which basic principles and advantages of self-assessment were discussed on the basis of students' knowledge from the methodology classes and their experience as learners. The students were then asked to design in groups a self-assessment form that could be used to evaluate their speeches. Most students agreed that such a form should include the following elements: topic, organization, pronunciation, grammar, vocabulary and general impression. The students were not provided with strict instructions as to what form their evaluation should take, other than a very general guideline: "Evaluate the speech, mention some strengths and weaknesses, comment on pronunciation, grammar, vocabulary, contents and organization". Most students found the tasks quite difficult despite the training in self-assessment they underwent in their first year. Most of them were not able or willing to identify their strengths and weaknesses and limited their evaluations to very general statements as in the examples below:

Student 1:

- the whole text is rather chaotically organized; I should focus on coherence more;
- the vocab is not sufficiently advanced;
- sometimes I tend to mispronounce the sounds in the end of the words;
- I have some problems with diphthongs e.g. like in 'follows'.

Student 2:

I believe that both the topic and the word choice are adequate, I did my best to be as fluent and understandable as possible.

Student 3:

I think that there is more advanced vocabulary in my speech. However, I still have to practice on my fluency during speaking. I don't see any major grammatical mistakes. In my opinion, the speech is very logical and interesting.

Student 4:

I might have made grammatical mistakes, some problems with pronunciation, I used vocabulary which I have learnt recently.

As was already said above, the students rarely mentioned their strengths and weaknesses; instead they often limited their evaluation to listing examples of mistakes they made and providing the correct versions, as in the examples below:

Others instead of 'other'.

Worries instead of 'worrying'.

Similar to instead of 'equally similar to'.

Not 'prohibit but pro'hibit.

Should be fight 'off' not fight 'down'.

Problems with pronouncing the word 'vulnerability'.

Some students provided very detailed assessments closely following the format discussed in class, as in the following example:

Topic: As I said in my speech, the subject of abortion will always be a popular and controversial topic for discussion. Since there are many various aspects mentioned by both the opponents and the supporters of abortion, I found the topic interesting to be discussed.

Organization: I think that my speech is clear and well organized. There was the introduction, the main body, in which I mentioned what others claim and my own opinion, and the conclusion.

Pronunciation: I have noticed some mistakes, for example in *religious*.

Vocabulary: I believe that the vocabulary and the expressions that I have used this time are more advanced than those that I used for my previous speech.

Grammar: I have not noticed any mistakes.

Impression: Generally, I am satisfied with both the fluency and accuracy. I find this recording better than the last one.

Finally, out of the thirty students who turned in their written evaluations, four designed their own forms in which they included sections such as: *What I like in my performance* or *What I don't like in my performance*, which could be considered to be variations of the strength/weaknesses categories, or lists of the new vocabulary items that the students deliberately tried to include in their recordings. They also used colors and plus and minus signs to indicate different aspects of their speech performance.

On the whole, the results of the first attempt at introducing self-assessment into the course were rather disappointing as most students were not able to evaluate their speeches effectively, despite the previous training they received. This observation led the authors to the conclusion that a simple self-assessment tool may provide a way of helping students focus on specific features of their speeches. The development and implementation of the checklist will be described in the next section.

4.4 Implementing a Self-Assessment Checklist

In creating the self-assessment checklist, we were inspired by several sources, namely: the *Common European framework of reference (CEFR)*, the guidelines for the practical English oral exam at the Teacher Training College in Poznań, as well as the authors' teaching experience and observations. The descriptors used in the *European language portfolio*, specifying the language level of learners, formed a basis for a detailed approach to the criteria incorporated in the preparation of the list. Since the graduates of the Teacher Training College are expected to reach the level specified as C1, the relevant description from the CEFR was taken into account, thus setting the frame within which the authors intended to operate. The guidelines for the final practical English exam, containing specific suggestions as to the assessment of pronunciation and the use of English, provided a substantial inspiration for the authors at the initial stage of the process. Some of the CEFR level

C1 descriptors, which guided the authors in their work, included the following (CEFR 2001, pp. 74–78):

- (a) “Can express himself/herself fluently or spontaneously, almost effortlessly. Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions. There is little obvious searching for expressions or avoidance strategies; only a conceptually difficult subject can hinder a natural smooth flow of language”.
- (b) “Can use language flexibly and effectively for social purposes, including emotional, allusive and joking language”.
- (c) “Can argue a formal position convincingly”.
- (d) “Can produce clear, smoothly flowing, well-structured speech, showing controlled use of organizational patterns and a wide range of cohesive devices”.

The guidelines for assessing students at the final practical English oral exam, focusing on the language and communication skills, were also taken into consideration, including the following (Regional practical English test specifications 2005):

- (a) the use of English: structure of sentences, the use of tenses, verb forms, the use of articles, collocations, advanced vocabulary;
- (b) pronunciation: the quality of vowels and consonants, intonation patterns, word stress, fluency;
- (c) communication skills: expressing personal opinions, asking and answering direct questions, interacting constructively.

The role of the last aspect listed in the guidelines, namely interacting constructively, was considered to be of marginal importance only, as students recorded their monologues rather than conversations (with two exceptions). The contribution or expressing opinions were crucial in the presentations, but not, for obvious reasons, ‘responding’ to each other’s comments.

4.4.1 The Checklist

As a result, the following list of criteria was created and sent out to the students once they finished their independent, unguided evaluation. There were three categories in the checklist, the first being the *evaluation of the content* (points one to six), the second dealing with *pronunciation* (points seven to twelve), and the third covering *vocabulary* (points thirteen to seventeen). Twenty-five students completed and submitted the self-assessment checklist (Table 1).

The list was preceded by the title, the assessment scale and the name slot. At the bottom of the table, space was provided for students where they could reflect on their strongest and weakest points, or add other comments they might have been willing to share with the authors. Below that section, the evaluation-scale of the checklist itself was added, so that students could circle the phrase they agreed with, choosing one of the following answers: “very useful”, “useful”, “not useful”, “not useful at all” and “cannot say”, or write about their reactions to the checklist in their own words.

Table 1 Self-assessment checklist

| | In my speech I managed to... | 5 | 4 | 3 | 2 | 1 | Comments/ examples |
|-----|---|---|---|---|---|---|-----------------------|
| 1. | Clearly present complex ideas | | | | | | |
| 2. | Adequately describe experience | | | | | | |
| 3. | Follow a logical order of events | | | | | | |
| 4. | Emphasize important arguments | | | | | | |
| 5. | Give appropriate examples | | | | | | |
| 6. | Draw conclusions | | | | | | |
| 7. | Speak fluently | | | | | | |
| 8. | Use appropriate intonation | | | | | | |
| 9. | Pronounce “th” correctly | | | | | | |
| 10. | Pronounce final voiced consonants correctly | | | | | | |
| 11. | Pronounce vowels correctly | | | | | | |
| 12. | Use correct stress in words | | | | | | |
| 13. | Use advanced vocabulary | | | | | | |
| 14. | Vary sentence structures | | | | | | |
| 15. | Form grammatically correct sentences | | | | | | |
| 16. | Use correct verb forms | | | | | | |
| 17. | Use correct articles | | | | | | |

4.4.2 Students’ Responses in the Checklist

When completing the checklist, the students used the grading scale from 1 to 5, 5 being the highest. Fifty percent of those who handed in their checklists seemed to be satisfied with their own performance, as they marked the first three columns, assessing different aspects of the presentation as “very good”, “good” or “quite good”. There were ten students who ticked the column with grade 2 when assessing some aspects of the speech, and three students who marked the last column (grade 1) when assessing their pronunciation and intonation. As many as fifteen students wrote additional comments in the column designed for that purpose. The extent to which they elaborated on a given aspect of their speech varied from just one simple sentence or phrase to a few sentences. The following remarks were included in the checklist:

- There were moments where I could have spoken clearer.
- I mispronounced some words.
- I sound more accurate than fluent.
- Unfortunately, this time it seems there are no conclusions at the end of my speech.
- My speech was quite fluent.

- I still have problems with my intonation.
- I haven't noticed incorrect verb forms.

Quite frequently, the students wrote the word where they noticed a mistake or a phrase which they seemed to be proud of:

- Recurring, to outline, to plunge into.
- I made a mistake in phenomena, where I should say phenomenon.

All the students used the space provided below the table and listed their strongest and weakest points, although there was one person who did not provide any example of a strong point and wrote "lack" in that line. However, in most cases, not only good but also poor aspects of the students' speeches were enumerated, with the focus on grammar, pronunciation and vocabulary:

- Poor grammar and vocabulary variety.
- I knew what I wanted to say but somehow I couldn't put my thoughts into words.
- Final devoicing was my major problem, and vowels were sometimes carelessly pronounced.
- I think the organization, argumentation and the presentation of the topic itself is my strong point.
- The last point below the checklist, namely the evaluation of the criteria, was completed by all students and, subsequently, discussed in the interview.

4.5 The Interview

The next stage of the action research included meeting with the students and talking about the self-assessment checklist, as well as reflecting on the process of self-assessment itself. In the time available, 15 students were interviewed and recorded. The questions asked during the interview focused primarily on the evaluation of the checklist, on its wording or structure specifically, but they were also concerned with the strategies students use to learn the language, and their previous experience in self-assessment. They were as follows:

- (1) How would you evaluate the checklist?
- (2) Which of the statements were unclear or difficult to understand?
- (3) Which of them would you change?
- (4) How do you intend to improve the aspects of speaking which you evaluated as poor?
- (5) What do you usually do in order to improve your language?
- (6) Do you use any of the strategies acquired in your first year learner training?
- (7) Did you assess yourself or record your speeches before you started studying at the College?

- (8) Do you think you might be willing to introduce self-assessment in schools when you become a teacher?

On average, the interview lasted 10–20 min and was arranged with each student on an individual basis.

4.5.1 Students' Responses to Interview Questions

With reference to the first question, the majority of students interviewed stated that the list of criteria was useful. One student considered it very useful, three stated that it was not useful and one was skeptical about the whole idea of self-assessment and marked the answer “cannot say” in the checklist. Among the reasons justifying the usefulness of the list was the fact that it was easier for students to assess their own speeches with the help of the checklist, while without it the task seemed much more difficult: “I like it when all the points are listed, because when we had to assess the speech I had problems what to write; we didn't have such a pattern, didn't know what to pay attention to”. Thus, some respondents stated that it would be ideal to have access to the list even before the recording, so that they could be aware what to focus on. Three students who stated that the list was not useful were actually expressing doubts about the “whole idea of self-assessment”. They questioned the process of having to record themselves, of listening again to their own speeches and then evaluating the presentation. Moreover, having to transcribe their presentations seemed to be too troublesome for them: “writing the transcript takes at least 2 h and it's horrible”. Although the whole idea of self-assessment was “useless”, as the three students frequently repeated in the interview, the checklist itself was evaluated more positively: “we know what to pay attention to; it's better than transcribing the whole speech”. The student who circled “cannot say” in the evaluation thought it would have been better to have been provided with the checklist at the beginning of the course, because “everyone had worked out the system” by the time the second recording was prepared.

On the other hand, however, there was a more positive response coming from a student who expressed a very enthusiastic view on the process of self-assessment: “I would convince my friends to listen to the recording at least four times, devote half an hour instead of 5 min—during my third listening I heard a lot more than during the first. It's simply impossible to focus on all aspects in one go: you need to listen again and again because each time you focus on a new aspect”. Another student admitted that although she understood why some people might dislike evaluating their speeches and the very process of recording, for her it was extremely useful, because when she heard herself, she could notice her mistakes: “it helped me a lot, because I paid attention to what I wanted to improve, for instance, I wanted to change intonation”. She also added: “many people are not aware of some aspects; such a checklist can help you focus on your weak points”. Additionally, the student emphasized that for her it was an exceptionally good idea to transcribe the speeches: “although I know it might be difficult, I think that only when I transcribed

whatever I said, was I able to notice details, and details are important because we are going to become teachers, so we should be aiming at perfection; I noticed some vowels then, or some devoicing, and I think it would be best to do both, transcription and the checklist, because we don't always hear the mistakes we make". It may be interesting to note here that both students prepared a detailed analysis of their speeches, with the checklist criteria discussed and fully described in the comments.

As far as the second question is concerned, six respondents considered the second and the sixth criterion to be difficult to understand, as "adequately describing experience" and "drawing conclusions" seemed to have been too vague for them. One student had problems with evaluating fluency ("how do I do that?"). There were also a few students who expressed doubts about the meaning of the criterion: "I managed to follow a logical order of events" or "I managed to emphasize important arguments". They all agreed, however, that those criteria would probably depend on the type of speech prepared, so they would not be relevant in all presentations. As far as the third question is concerned, there were suggestions from two students that perhaps more specific descriptions should be added, as, for instance, "th should be characterized more fully", or more attention should be paid to the pronunciation of those segments which are "difficult for Poles, some diphthongs, for example".

Answers to questions seven and eight show that none of the students interviewed had any practice in self-assessment before entering university. Most students agreed, however, that it might be a good idea to introduce self-assessment in their classroom once they become teachers: "I think everyone should evaluate themselves; everyone should be able to say what their weak aspect is; we often think that teachers and friends exaggerate when they say that something is wrong, but when you listen to yourself, you will hear they are right". The three students who disapproved of self-assessment in general expressed a negative attitude towards using this strategy in their future teaching practice, arguing that it would be too difficult for their pupils: "why should a student in high school know how to assess himself? They don't learn pronunciation at school, or transcription".

The questions which referred to the students' ways of learning and improving the language, namely numbers four, five and six, triggered a variety of responses. Among the strategies used by the students the following were quoted: learning useful words, using word cards, drawing trees, mnemonic techniques, as well as listening to songs and watching films. It is interesting to note here that those ways of learning had been developed before their first year at the college and the course in study skills helped them become more aware of their activities ("I appreciate that more now, use it more consciously; before I came here I didn't know it was a technique; I thought everybody had to cope somehow; I didn't know others had the same problem and that there was a theory about it"). When inquired further about the source of help in acquiring new strategies, most students admitted: "Nobody helped me, I found them on my own". One student found inspiration in the preparation for the final exam in senior high school, and was directed and guided by a teacher.

When asked for a comment on the ways of learning, the three students who disapproved of self-assessment stated: “we all have our own ways of learning, difficult to change” and, in response to the fourth question (“How do you want to improve the aspects of language which you evaluated as poor?”), a surprising reply was given: “I think we should have pronunciation in the third year”. Moreover, the question “What can you yourself do?” triggered a reply in the past tense: “I was listening to longer speeches after the first year exam”, signaling to the interviewer that, in fact, the ways of learning the language have not changed or improved since then.

5 Conclusions

Undoubtedly, there are many aspects of the interviews and the checklist which could be further analyzed and studied, but the most crucial goal was attained: we now have learnt that the majority of students interviewed found the checklist useful and would consider using self-assessment in the classroom once they become qualified teachers. They would also prefer to have been given the checklist earlier, rather than later in the academic year, as evaluating themselves on their own turned out to be very difficult for many. A few points in the list of criteria should be improved, or made more precise, such as fluency, describing experience, or using advanced vocabulary. The study revealed that it is necessary to raise students’ awareness about the importance of self-assessment in language learning as well as to provide them with systematic training in this skill. Self-assessment is a difficult process and it is influenced both by the ability of the students as well as their attitude to it. Some, more autonomous and independent students, will be eager to experiment with it and will be able to design their own techniques of conducting it. More teacher dependent students will need more time to be convinced about the usefulness of self-assessment and will need more guidance before they are willing to experiment with it. It seems that active involvement of the students in the process of designing the instrument is crucial if students are to accept it as their own. The students should also be allowed to experiment with different assessment instruments and choose aspects of the target language performance they want to focus on at any given time. Finally, the issue of the accuracy of students’ self-assessment and the correlation between students’ and teachers’ assessment should be addressed.

References

- Alderson, J. C. and L. F. Bachman. 2004. Series editors’ preface to *Assessing speaking*. In *Assessing speaking*, S. Luoma, ix–xi. Cambridge: Cambridge University Press.
- Brown, H. D. 2001. *Teaching by principles: An interactive approach to language pedagogy*. (second edition). White Plains, NY: Pearson Education.
- Brown, H. D. 2004. *Language assessment: Principles and classroom practices*. Harlow: Pearson Education.

- Brown, G. and G. Yule. 1983. *Teaching the spoken language: An approach based on the analysis of conversational English*. Cambridge: Cambridge University Press.
- Bygate, M. 1987. *Speaking*. Oxford: Oxford University Press.
- Bygate, M. 1996. Effects of task repetition: appraising the developing language of learners. In *Challenges and change in language teaching*, eds. Willis, J. and D. Willis, 134–146. London: Heinemann.
- Bygate, M. 1999. Tasks as context for the framing, reframing and unframing of language. *System* 27: 33–48.
- Carr, W. and S. Kemmis. 1986. *Becoming critical: Education, knowledge and action research*. London: Falmer.
- Council of Europe. 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Hughes, R. 2011. *Teaching and researching speaking*. Harlow: Pearson Education.
- Johnson, K. 2008. *An introduction to foreign language learning and teaching*. Harlow: Pearson Longman.
- Kemmis, S. and R. McTaggart. 1988. *The action research reader*. (third edition). Victoria: Deakin University Press.
- Levelt, W. J. M. 1989. *Speaking: From intention to articulation*. Cambridge: MIT Press.
- Luoma, S. 2004. *Assessing speaking*. Cambridge: Cambridge University Press.
- Nunan, D. 1989. *Understanding language classrooms. A guide for teacher-initiated action*. Englewood Cliffs, NJ: Prentice Hall International.
- Regional Practical English Test Specifications: Revised and updated*. 2005. Poznań: Kolegium Języków Obcych UAM.
- Skehan, P. 1998. *A Cognitive approach to language learning*. Oxford: Oxford University Press.
- Thornbury, S. 2005. *How to teach speaking*. Harlow: Pearson Education.
- Zimmerman, B. J. 2000. Attaining self-regulation: A social cognitive perspective. In *Handbook of self-regulation*, eds. M. Boekaerts, P. R. Pintrich and M. Zeidner, 14–35. San Diego, CA: Academic Press.