UvA-DARE (Digital Academic Repository)

The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency

Hulstijn, J.H.

Publication date
2007

Published in
The Modern Language Journal

[Link to publication](#)

Citation for published version (APA):

Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91, 662-666.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Contribution to *The Modern Language Journal*, Vol 91, as an invited comment on David Little's position paper "The Common European Framework of Reference for Languages: Perspectives on the making of supranational language education policy".

Word count: 3238 words text, excluding title, including references.

The shaky ground beneath the CEFR:
Quantitative and qualitative dimensions
of language proficiency¹

Jan H. Hulstijn, Amsterdam Center for Language and Communication, University of
Amsterdam

The *Common European Framework of Reference for Languages* (CEFR, Council of Europe, 2001) has proved to be extremely influential in the promotion of plurilingualism in Europe, in syllabus design, curriculum planning, and in language examinations in a number of European countries. This is a welcome trend that the many language experts, educational officers and politicians who created, designed, promoted, and implemented the framework should be congratulated on. Language learners, language teachers, educational institutions and employers will find the framework a helpful tool in the setting of curricular goals and entry requirements, in comparisons of curricular systems in various countries and regions, and in communicating in rather concrete terms about what language learners can and cannot do in their foreign language(s). The framework is, of course, not perfect but good enough to be improved (Alderson, Figueras, Kuijper, Nold, Takala & Tardieu, 2004) and developed further (The Council of Europe held a forum in Strasbourg in February 2007 to consider broadening and deepening the framework). As Little (this volume) observes, the success of the CEFR might well be caused by its combination of what is familiar (the traditional distinction between 'beginner', 'intermediate', and 'advanced' levels) and what is new (an elaborate system of descriptors giving communicative content to the levels beginner/basic, intermediate/independent, and advanced/proficient).

¹ I am most grateful to John de Jong (Pearson Education) and Charles Alderson (Lancaster University) for their critical comments on an earlier version of this text.

Having praised the CEFR enterprise and having committed myself to its goals in the opening lines, I will devote the remainder of this contribution to a presentation and discussion of some abstract, theoretical issues, underlying the CEFR.

The notion of language proficiency presented in the CEFR rests on two pillars, quantity and quality, closely intertwined. Proficiency descriptors, such as the example of Overall Oral Production at the B2 level (CEFR 2001, p. 27, 58) cited here, show how elements of quantity (what the learner is able to do) and elements of quality (how well the learner is able to do this) are interwoven.

B2. Can give clear, detailed descriptions on a wide range of subjects, related to his/her field of interest, presenting it as a linear sequence of points.

The element “descriptions on a wide range of subjects, related to his/her field of interest” indicate *what* the learner can do, while the elements “clear, detailed”, “presenting it as a linear sequence of points” indicate *how well* the learner can do this. According to De Jong (2004), “quantity refers to the number of domains, functions, notions, situations, locations, topics, and roles that a language user can deal with. The notion of quality refers to (1) the degree to which language use is effective, leading to degree of precision, in understanding what is meant, and in expressing one’s meaning, and (2) the degree to which language use is efficient, leading to communication with least possible effort”²

The CEFR not only provides scales of the mixed *What & How well* type (p. 58 – 96) but also scales of a number of “linguistic competences” (p. 108 – 118), such as Vocabulary Range, Vocabulary Control, Grammatical Accuracy, and Phonological Control. The question then arises whether it is necessarily true that a learner, who is placed at the B2 level of overall production cited above, must also have attained the B2 level of all the linguistic competences scales, or whether it is possible for a learner to be situated at different levels of different

² Quotation, with permission, from De Jong’s powerpoint presentation.

scales. In principle, one could conceive of three types of L2 users: (1) people who can do only few things in terms of quantity but their performance is characterized by high linguistic quality, (2) people who can do many things in terms of quantity but their performance is characterized by low linguistic quality, and (3) people whose quantity range matches their performance quality, as suggested by the CEF scales of the mixed type.³ This is an empirical issue which deserves to be properly investigated. In a project called “What is speaking proficiency?” (WiSP), conducted at the Amsterdam Center for Language and Communication (ACLC) of the University of Amsterdam, we are currently investigating what variability we find among 200 adult nonnative speakers of Dutch, who, in terms of functional adequacy, perform at the B1 or C1 levels of speaking proficiency, in their control of a number of linguistic subskills, and what variability we find among a control group of 50 native speakers (see www.hum.uva.nl/wisp).

The issue of the association of quantity and quality of language proficiency leads to some truly fundamental questions concerning the relationship between forms and function (meaning) in language (Newmeyer, 1998). One fundamental question is whether a sharp borderline can be drawn between linguistic and nonlinguistic cognition. In favor of the formalist view, I would argue that there are elements of language that are not, or hardly, related to meaning. Some grammatical properties are not concerned with meaning, such as the positions of object and verb relative to each other (VO versus OV, in languages such as English and Japanese respectively) and the absence or presence of certain phonemes. Other elements of language are clearly related to meaning. This is evidenced by the word category of content nouns. Knowledge of English common nouns such as *eat* and *kitchen* and proper nouns such as *Berlin* includes knowledge of their meaning, showing that linguistic cognition is associated with nonlinguistic cognition (knowledge of the world).

³ In fact, the CEFR allows an almost infinite number of profiles. I return to this point in the final section of this paper.

The authors of the CEFR are, of course, aware of these underlying fundamental linguistic, psycholinguistic and philosophical issues. But, as North and Schneider (1998, p. 242) correctly observe, the CEFR authors, in the absence of fully developed and properly tested theories of language proficiency, had to go ahead by scaling the available descriptors. The task to develop and test a theory of language proficiency, however, still remains on our agenda. Many language-policy makers and educational professionals seem to ignore the need to conduct such research, pushing for further implementation of the CEFR and developing similar scales for other target groups. In the following section, I outline the beginnings of a theory of core language proficiency. I acknowledge that the theory pertains primarily to native speakers but we need to know first what language proficiency means in the case of native speakers before we can consider the case of nonnative speakers. Later on in this paper, I will consider the theory's implications for nonnatives and the use of the CEFR for L2 assessment.

Core language proficiency

As of today, my theory of core language proficiency, still in its infancy, can be formulated in three points.

1. I propose a global distinction between lower-order and higher-order cognition in language processing. *Lower-order linguistic cognition* pertains to the largely implicit, unconscious knowledge of speech sounds, prosody, phonological and syntactic knowledge of a number of lexical items, morphology and syntax, and the automaticity with which this knowledge can be processed. *Higher-order cognition* pertains to largely, but not exclusively, explicit, conscious knowledge of all sorts of topics (of which vocabulary size is but one rough index), attention allocation, decision making, inferencing ability, and the like. *Working-memory capacity* is taken as a potentially important moderator variable in the online use of lower and higher-order cognition. Empirical support for the distinction between lower-order and language-

specific cognition on the one hand, and higher-order and non language-specific cognition on the other hand, in L2 learners was produced in several studies, conducted at the University of Amsterdam (Schoonen et al., 1998, 2003; Van Gelderen et al, 2003, 2004).

2. I hypothesize that all humans, even adults of older age and lower educational background, are capable of implicit learning, i.e., capable of acquiring lower-order cognition in all walks of life, in an implicit way (Ellis, 2005; Hasher & Zacks, 1979; Hulstijn, 2002, 2003, 2005; Paradis, 2004; Reber, Allen & Reber, 1999; Van Gerven et al., 2002). Implicit learning applies also to the acquisition of oral language skills (listening and speaking), in a second language (L2) just as well as in a first language (L1).

3. I propose the notion of *core language proficiency*. It differentiates between knowledge and skill. In the case of speech perception (listening) and speech production (speaking), *knowledge* refers to the mental representation of (1) speech sounds, phonemes, stress and intonation patterns, (2) lexical items, and (3) morphology and syntax.

Skill refers to the ability to accurately online process phonetic, lexical and grammatical information receptively and productively. The core of language proficiency restricts this knowledge and skill to frequent lexical items and frequent grammatical constructions, that is, to lexical items and syntactic constructions that may occur in any communicative situation, common to all adult native speakers regardless of age, educational level or literacy.

From this definition of core language proficiency, the claim can be derived that all (mentally healthy) adult native speakers, notwithstanding differences in age and intellectual functioning, are able to correctly and quickly understand (segment) and comprehend isolated utterances consisting of high-frequency lexical phrases and high-frequency morpho-syntactic structures, when perceived under normal acoustical conditions. The extent to which this provocative claim is correct with respect to receptive oral language processing (listening), will be investigated in a new project (2007-2011), conducted at the ACLC, with 200 adult natives

speakers, differing in age (20-65) and educational background (lower-level vocational training versus college level), and 100 nonnative speakers of Dutch.

The definition of core language proficiency does not show where exactly the borderline between core and periphery is to be found. I suggest that the demarcation line must be seen in probabilistic terms. On the basis of statistical analyses conducted on corpora of a wide variety of oral and written discourse, it should be possible to define the set of words and multiple-word constructions that have a high probability of occurring in communicative situations that all adult native speakers, regardless of age and educational level, are likely to be able to deal with. In our new project, three of these parameters (utterance length, grammatical complexity, and lexical difficulty) will be manipulated in order to test to what extent the notion of 'core language proficiency' can be empirically upheld.

Balance and imbalance of the quantity and quality dimensions of language proficiency

Let us return to the L2 user typology given above: (1) L2 users with a narrow range in terms of quantity but great depth in terms of linguistic quality, (2) people with a broad quantity range but little linguistic quality, and (3) people, as defined in the CEFR, whose quantity range matches their performance quality. Do these types, conceivable in principle, really exist? On the basis of the embryonic theory of core language proficiency in the domains of receptive and productive use of oral language, outlined above, I hypothesize that adult native speakers share a core of linguistic knowledge and psycholinguistic skills. They make use of this mainly implicit cognition in all oral communicative situations and do so fluently and automatically. Thus, as far as proficiency in these communicative situations is concerned, I hypothesize that adult native speakers with lower educational background are of type 1. In other words, the core makes up most of the language proficiency of natives speakers with low

educational background; the part of their proficiency beyond the core is relatively small. Native speakers with higher educational background are indistinguishable from natives at lower educational levels with respect to core language proficiency, but their profiles will differentiate beyond the core. The part of their proficiencies beyond the core is relatively large. Beyond the core, in principle, all three user types should be found.

For nonnative speakers, profiles may be expected to be different from profiles of native speakers. With respect to situations of oral communication that all native adults may find themselves in, imbalances are likely to emerge in the case of L2 learners with little exposure to and practice in using L2. They may have to compensate for a lack of linguistic knowledge and skill, especially in the lexical domain, with non-verbal strategies to achieve their communicative goals. This imbalance, with quality lagging behind quantity, is typical of the levels A1, A2 and B1. The CEFR scales reflect this situation. However, just as there is no empirical evidence (yet) for my hypotheses, the CEFR scales lack empirical support of what L2 specific knowledge and skill is minimally required for performance considered adequate in terms of communicative functioning.

The urgent need for empirical studies of L2 and L1 users

The CEFR enterprise is good for Europe and its citizens. However, we must not forget that its empirical base consists of judgments of language teachers and other experts with respect to *the scaling of descriptors*. In some cases, teachers had to take one of their students as a point of reference, judging whether that student was or was not able to perform what was stated in the descriptors. It is crucial to note, however, that the CEFR is not based on empirical evidence taken from *L2-learner data*. Psychometrically speaking, North and Schneider (1998, p. 238-239) are correct when they claim that the CEFR scales can be seen as unidimensional.

There is no, or insufficient evidence for unidimensionality of the notion of language proficiency, however, when it comes to *empirical studies conducted with L2 learners*.

First, to my knowledge, there are no longitudinal studies showing that all L2 learners at some functional level other than A1 (e.g. B2), arrived at that level by passing the level below (B1, in this example). In other words, there is no empirical evidence that, in Overall Oral Proficiency (p. 58), all learners first attain the functional level of A1, then the level of A2, etc., until they reach their individual plateau. Second, when we look at the CEFR from the L2 user's perspective, there is no empirical evidence showing that all L2 learners at a given level (other than the lowest level A1), are able to perform all tasks associated with lower levels (apart from occasional performance lapses), which should be the case if the CEFR scales are genuinely implicational and unidimensional. Third, and more seriously, there is no evidence in terms of learner performance that a learner at a given level of an overall scale (e.g. B2 Overall Oral Production) necessarily possesses the quality in terms of the *linguistic* scales at the same level (e.g. B2 Vocabulary Range, B2 Grammatical Accuracy, and B2 Phonological Control). I acknowledge that the CEFR document (2001) nowhere explicitly asserts or claims this to be the case. However, by distinguishing six levels on each scale and by using the same set of symbols to refer to these levels (A1, A2, B1, B2, C1 and C2), the document at least leaves the impression of suggesting that the normal match between functional and linguistic skills holds at levels indicated with the same symbol.

The three observations in the previous paragraph demonstrate the urgent need to empirically test the implications of the CEFR using real L2 learners rather than teachers rank ordering descriptors (see also Alderson, this issue). Theories of language proficiency must be developed and tested. (An embryonic example of such a theory was outlined above.) Research on developmental routes in second language acquisition (SLA), must be linked with language assessment research. Furthermore, there is a great need for corpus research. On the basis of

statistical analyses conducted on corpora of a wide variety of oral and written discourse, it should be possible to define the set of words and multiple-word constructions that have a high probability of occurring in certain communicative situations, especially in situations that all adults are likely to be able to deal with. It is high time that researchers of SLA, researchers of language assessment, and corpus linguists paid attention to each other's work and engage in collaborative research, testing the linguistic, psycholinguistic and sociolinguistic assumptions on which the CEFR rests.

In the Netherlands, all buildings are built upon poles, invisible to the eyes but of truly "fundamental" importance. To date, the CEFR only rests on the poles of teacher perceptions. Valid and reliable as they are or may be, they provide a foundation too weak for the CEFR building with its heavy-weight implications for language-education policy in Europe. Educational authorities and politicians must be made aware of the missing linguistic, psycholinguistic and sociolinguistic poles underneath the CEFR and the urgency of making funds available for collaborative research. To provide a building with a proper foundation, it need not be torn down first and built up later. The educational community can continue to reside in the CEFR, while researchers are constructing the poles underneath.

References

- Alderson, J. C., Figueras, N., Kuijper, H., Nold, G., Takala, S. & Tardieu, S. (2004). *The development of specifications for item development and classification within the Common European Framework of Reference for Languages : Learning, teaching, assessment. Reading and listening. Final report of the Dutch DEF Construct Project*. Available on request from J. C. Alderson, <c.alderson@lancaster.ac.uk>.
- Council of Europe. *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press / Council of Europe, 2001.

- De Jong, J.H.A.L. (2004). Comparing the psycholinguistic and the communicative paradigm of language proficiency. Presentation given at the international workshop "Psycholinguistic and psychometric aspects of language assessment in the Common European Framework of Reference for Languages". University of Amsterdam, 13-14 February, 2004.
- Ellis, N.C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Languages Acquisition*, 27, 305-352.
- Hasher, L., & Zacks, R.T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.
- Hulstijn, J.H. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, 18, 193-223.
- Hulstijn, J.H. (2003). Incidental and intentional learning. In C. Doughty & M.H. Long (Eds.), *Handbook of Second Language Research* (pp. 349-381). London: Blackwell.
- Hulstijn, J.H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition*, 27, 129-140.
- Little, D. (2007). The Common European Framework of Reference for Languages: Perspectives on the making of supranational language education policy. *The Modern Language Journal* (this issue).
- Newmeyer, F.J. (1998). *Language form and language function*. Cambridge, Mass.: MIT Press.
- North, B., and Schneider, G. (1998). Scaling descriptors for language proficiency scales. *Language Testing*, 15, 217-262.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: Benjamins.
- Reber, A.S., Allen, R., & Reber, P.J. (1999). Implicit versus explicit learning. In R.J. Sternberg, (Ed.), *The nature of cognition* (pp. 475-513). Cambridge, MA: MIT Press.

- Schoonen, R., Hulstijn, J., & Bossers, B. (1998). Metacognitive and language-specific knowledge in native and foreign language reading comprehension: An empirical study among Dutch students in grades 6, 8 and 10. *Language Learning*, 48, 71-106.
- Schoonen, R., Van Gelderen, A., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P. & Stevenson, M. (2003). First language and second language writing: the role of linguistic fluency, linguistic knowledge and metacognitive knowledge. *Language Learning*, 53, 165-202.
- Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Snellings, P., Simis, A. & Stevenson, M. (2003) Roles of linguistic knowledge, metacognitive knowledge and processing speed in L3, L2 and L1 reading comprehension: A structural equation modeling approach. *International Journal of Bilingualism*, 7, 7-25.
- Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first- and second-language reading comprehension: A componential analysis. *Journal of Educational Psychology*, 96, 19-30.
- Van Gerven, P. W. M., Paas, F. G. W. C., Van Merriënboer, J. J. G. , & Schmidt, H. G. (2002). Cognitive load theory and aging: effects of worked examples on training efficiency. *Learning and Instruction*, 12, 87-105.