

# Linguistic Skills of Adult Native Speakers, as a Function of Age and Level of Education

---

<sup>1</sup>KIMBERLEY MULDER and <sup>2,\*</sup>JAN H. HULSTIJN

<sup>1</sup>Max Planck Institute for Psycholinguistics, Nijmegen and <sup>2</sup>University of Amsterdam  
\*E-mail: j.h.hulstijn@uva.nl

This study assessed, in a sample of 98 adult native speakers of Dutch, how their lexical skills and their speaking proficiency varied as a function of their age and level of education and profession (EP). Participants, categorized in terms of their age (18–35, 36–50, and 51–76 years old) and the level of their EP (low versus high), were tested on their lexical knowledge, lexical fluency, and lexical memory, and they performed four speaking tasks, differing in genre and formality. Speaking performance was rated in terms of communicative adequacy and in terms of number of words, number of T-units, words per T-unit, content words per T-unit, hesitations per T-unit, and grammatical errors per T-unit. Increasing age affected lexical knowledge positively but lexical fluency and memory negatively. High EP positively affected lexical knowledge and memory but EP did not affect lexical fluency. Communicative adequacy of the responses in the speaking tasks was positively affected by high EP but was not affected by age. It is concluded that, given the large variability in native speakers' language knowledge and skills, studies investigating the question of whether second-language learners can reach native levels of proficiency, should take native-speaker variability into account.

## INTRODUCTION

The notion of *native speaker* (NS) is relevant in much of the second-language (L2) acquisition literature. For instance, one of the most fundamental issues is concerned with the question of whether it is possible for L2 learners to attain native, or near-native, levels of L2 proficiency. In empirical studies, investigating this issue, the notion of native or near-native level of proficiency has to be operationalized. The practical question then is: What counts as NS proficiency and NS performance? A related question is concerned with the thorny issue of how to define balanced bilingualism. Although a balanced bilingual is not necessarily two monolinguals united in a single person, in empirical studies bilinguals are usually compared with monolingual native speakers (for a recent review of the literature, see De Groot 2011: 4–5). Thus, for the study of at least two fundamental questions it is important to define what a native speaker is.

Escudero and Sharwood Smith (2001: 284) propose that the notion of prototypes can be fruitfully used to conceptualize the notion of NS (see also Davies 2003). According to the authors, a linguist may have the following view of a

NS. In extralinguistic terms, the prototypical NS acquires the language in an 'initial language environment (age 2–5)', maintaining the language into adulthood. In intralinguistic terms, the prototypical NS is characterized by a grammar, basic lexicon, and accent. The authors place paralinguistics and orthography in the periphery, while characterizing grammatical intuitions, pragmatics and extended lexicon as 'possibly prototypical, possibly peripheral', expressing the hope that 'future research may lead to refinements of the notion of linguistic ability in native speakers' (p. 285). Our study aims at contributing to achieving this goal by examining, in a sample of 98 adult native speakers of Dutch, the effects of age and level of education and profession (EP) in their performance in a range of tasks, tapping linguistic knowledge skills, linguistic processing skills, and communicative success in conveying the message in several speaking tasks. Before we present our research questions, we give an overview of previous research, mainly in the domain of psycholinguistics.

## Review of the literature

There are not many empirical studies examining how linguistic ability in native speakers (NSs) differs as a function of age (i.e. younger and older adults) or socio-economic status (SES). Most studies do so in the context of studying individuals with some type of language-related syndrome or impairment in comparison with unimpaired people, or in the context of educational issues, such as literacy. We begin our review with studies examining age effects, then review studies examining effects of level of education, and finally look at studies examining the interaction of age and education. As will become apparent below, most empirical studies focus on either skills that are seen as forms of declarative, crystallized knowledge or skills seen as the ability to rapidly process linguistic information. For ease of reference, we will refer to these two types as *knowledge skills* and (*speed of*) *processing skills*, representing, respectively, knowledge and processing components (and hence facets) of speaking proficiency. Space limitations prohibit an extensive review. For more information, we refer the reader to the following updates on the literature: Birren and Schaie (2006), Craik and Salthouse (2008) and De Bot and Schrauf (2009).

### Age effects

Virtually all studies examining age effects in NSs are concerned with the speed with which older people *process* linguistic information rather than with increases or declines in declarative *knowledge*. Many studies observed that people older than 50 process linguistic information more poorly than young adults, which may be caused by slowed processing speed, shrinking working memory, inhibitory deficits, and/or declining sensory function, such as decline in hearing capacities (Tun and Wingfield 1999; Burke *et al.* 2000; Schneider,

*et al.* 2002a; Waters and Caplan 2005; Schrauf 2008). According to Schrauf (2008), it is in tasks involving effortful rather than automatic processing that older adults show decrements relative to younger adults. Retrieval of the meaning of a word as a result of reading or hearing the word is an automatic process, in contrast to word finding, which requires an explicit search through the lexicon. More automatic mental processes are thus not as heavily affected by age as are more explicit mental processes.

### Level of education

In a study involving 1855 native speakers (age 24–81 years) Van der Elst *et al.* (2005) found that higher educated individuals perform better in word learning than lower educated individuals over the entire age range tested. Studies using picture description tasks observed a clear effect of education in terms of response length and completeness of the description (Beland *et al.* 1993; Le Dorze and Bedard 1998; Mackenzie 2000), with well-educated adults scoring higher on these tasks than their less well-educated peers (Hawkins and Bender 2002).

### Interactions of age and level of education

Higher education has been shown to ameliorate age-related cognitive decline (Bosma *et al.* 2003). On the basis of the findings in a word-learning study involving 338 participants of lower and higher education from four age groups (ranging from 24 till 76 years old), Meijer *et al.* (2008) argue that age-related changes in verbal learning may be smaller in higher educated individuals. The authors conclude that education appears to moderate verbal learning in later adulthood. Moreover, education also influences the ability to use longer inter-stimulus intervals to enhance performance in such a way that lower educated people need more study time to achieve the same level of performance as higher educated people, which suggests that lower educated people need more time to memorize verbal information in daily life. Using data from a picture description task, performed by 225 non-brain-damaged adults, Mackenzie *et al.* (2007) found that the number of concepts used accurately and completely, as well as the amount of topic subdivision, increased with amount of education. Patricacou *et al.* (2007) tested naming performance on the Greek version of the Boston Naming Task (BNT) in 100 native speakers of Greek, divided over four age groups and three levels of education. They observed strong main effects of age and education on naming performance. Moreover, there was a significant interaction between age, education, and gender. Thus, the studies reviewed in this section converge on the view that level of education mediates the negative effects of ageing in that people with higher education are less affected by information processing and memory problems associated with old age than are people with lower education.

## Design of the study and research questions

To empirically explore the notion of NS, as recommended by Escudero and Sharwood Smith (2001), we first investigated to what extent, in a sample of 98 adult NSs of Dutch, differences in their age and in the level of their EP are associated with their lexical knowledge, lexical fluency, and lexical working memory. Lexical knowledge was assessed with a vocabulary and a word association test, lexical fluency was assessed in four computer-administered speed tasks (reaction times), and lexical working-memory capacity with two span tests. The decision to tap these three types of linguistic cognition was based on the common distinction in cognitive psychology between, on the one hand, declarative, crystallized knowledge, and, on the other hand, the ability to rapidly process linguistic information and working memory (WM) (Anderson 1980; Paradis 2004; Ullman 2004). WM capacity in the verbal domain has been shown to be associated with performance in functional reading, listening, speaking, and writing tasks (Conway *et al.* 2007).

RQ1. To what extent do adult NSs, not impaired by mental disorders, differ in lexical fluency, lexical knowledge, and lexical working memory capacity, as a function of their age and level of EP?

While the first research question examined how NSs might differ in three subskills in the lexical domain, the remainder of the study included an examination of how NSs may differ in performing everyday speaking tasks. Participants performed four speaking tasks, differing in genre and formality.

RQ2. To what extent does the communicative adequacy with which adult NSs perform everyday speaking tasks differ as a function of their age and level of EP?

RQ3. To what extent do the length, lexical richness, fluency, and grammatical accuracy (grammatical errors) in the speaking-task performances differ as a function of participants' age and level of EP?

Linking the data collected in the first part of the study with those collected in the speaking tasks, we then aimed at answering the following research question.

RQ4. To what extent is the quality of the information conveyed (by adult NSs) in the responses in four speaking tasks, associated with participants' lexical skills (lexical knowledge, lexical fluency, and lexical working memory capacity), assessed independently from speaking?

Finally, in an attempt to gain a more complete picture of the way in which NSs might differ in their ability to get their message across in everyday speaking tasks, we examined how this ability might be affected by age and level of education on the one hand, and participants' lexical skills (processing, knowledge and memory) on the other hand.

RQ5. What is the relative contribution of participants' age, level of EP, and lexical skills (lexical knowledge, lexical fluency, and lexical working memory)

in the explanation of differences in the success with which adult native speakers convey the message in everyday speaking tasks?

## METHODS

### Participants

A total of 98 adult native speakers of Dutch, with normal or normal-to-corrected vision and hearing, living in and around Amsterdam, gave their written consent to volunteer as participant in the study. Given the fact that age-related declines in cognitive functioning begin relatively early in adulthood, reflecting a gradual decline from age 18 onwards rather than a sudden decline at a particular age (Salthouse 2009), we attempted to acquire participants in the 18–75 age range. In analyses of variance, they were arbitrarily divided into three age groups: 18–35 years old ( $n=42$ ), 36–50 years old ( $n=20$ ), and 51–76 years old ( $n=36$ ). To compensate for the disadvantages of using discrete, adjacent, arbitrary age groups, we conducted additional analyses using participants' real age (as a covariate or in the computation of Pearson  $r$  between age and other variables).

Participants also differed in educational and professional background. Educational level higher than or equal to *havo* (higher level diploma of secondary school) was classified as higher, while level of education lower than *havo* or equal to *mbo* was classified as lower. For the classification of the participants' professions, the SBC 92 classification list of the Statistics Netherlands (Centraal Bureau voor de Statistiek 2001) was used, distinguishing five levels of profession. Participants with levels 1–2 were coded as low, participants with levels 3–5 as high. The scores of both the EP levels were combined in the following way.<sup>1</sup> Individuals with low rankings on both education and profession (low–low) were assigned to the low EP group ( $n=52$ ), while all others (low–high, high–low, high–high) were assigned to the high EP group ( $n=46$ ). Table 1 gives the resulting partition of age and EP.

Table 1: Participant numbers by age and education–profession (EP)

Age group (years)	Low EP	High EP	Total
Young (18–35)	18	24	42
Older (36–50)	10	10	20
Senior (51–76)	24	12	36
Total	52	46	98

## Tasks and materials

Participants performed seven lexical tasks and four speaking tasks, administered in two to three sessions, totalling ~180 min. Three aspects of lexical skills (fluency, knowledge, and memory) were assessed, each aspect with at least two measures. *Lexical fluency* was tapped with four reaction-time measures: word association, auditory lexical decision, visual lexical decision, and picture naming. *Lexical knowledge* was assessed with two measures: accuracy scores in a paper-and-pencil vocabulary knowledge task and accuracy scores in the same word association task as used to assess lexical fluency. *Lexical memory* was assessed with an auditory and a visual word-span task.

### Word association task

This task was developed by M. Cremer and R. Schoonen (manuscript under review) and kindly made available to us. Each of the 69 trials of this task visually presented a prime word and two possibly associated words (a distracter word and a target word). Participants had to decide which of these words had the strongest relation to the prime word, by pushing a button corresponding to one of these words. The primes, target words, and distracter words were all high-frequency content words (nouns, verbs, adjectives, and one adverb), taken from a corpus of Dutch words addressed to young children (Schrooten and Vermeer 1994). The target words had a strong relation to the stimulus word based either on hyponymy (flower—tulip), hypernymy (apple—fruit), synonymy (steal—rob), meronymy (elephant—tusk) or they are a defining characteristic of the stimulus word (teeth—jaw). In contrast, the distracter words had a less strong relationship to the stimulus word: they did not relate in the ways mentioned above but bore an indirect or subjective relationship to the stimulus word. For example, an item could contain the prime *banaan* ‘banana’, the target word *fruit* ‘fruit’, which has a hypernymic relationship to the prime, and the distracter word *aap* ‘monkey’, which is less strongly related to the prime. The task was computer administered and controlled with the software package E Prime (Schneider *et al.* 2002b). We coded the speed (in ms) with which the response was given (as one of the four fluency variables) but also response accuracy (as one of the two lexical knowledge variables). The maximum accuracy score was 69; test reliability ( $\alpha$ ) was .87.

### Auditory and visual lexical decision tasks

For these tasks, we used the stimuli of the auditory and visual lexical decision tasks constructed by Poelmans (2003). Each task consisted of 50 existing Dutch words (mono-morphemic nouns) and 50 pseudo words. All items were highly frequent, belonging to the 2,143 most frequent Dutch words as listed by *Basiswoordenboek Nederlands* (De Kleijn and Nieuwborg 1996). To allow for a cross-modal comparison of the lexical decision tasks, half of the stimuli in the

auditory lexical decision task were used in the visual lexical decision task; the other half of the stimuli were taken from the visual lexical decision task used in the study of Poelmans (2003), also chosen from the *Basiswoordenboek Nederlands* (De Kleijn and Nieuwborg 1996). Both the word and non-word stimuli consisted of one, two, or three syllables.

### Picture naming task

The task used in this study is the first part of the picture naming task of De Jong *et al.* (in press), in which participants had to name 28 pictures, referring to common objects, as fast as possible. The task was controlled by E-prime and speech was recorded on a digital recorder.

### Vocabulary knowledge task

The task used in this study is a part of the vocabulary test of De Jong *et al.* (in press). Each item consisted of a sentence in which one word had been omitted and replaced by a visual gap and the first letter(s) of the missing word. For each gap, participants had to fill in a semantically appropriate word. Of the 120 items of the original test, the 60 most difficult items were selected for the present study.

### Auditory and visual word span tasks

In these tasks, participants listened to, or saw (respectively), sequences of Dutch common monosyllabic nouns, memorized them, and then wrote them down on a sheet of paper. The inter-word interval was 1,000 ms. The sequences increased gradually in length from two to eight words and remembering a sequence thus became more and more difficult. The maximum score was 8.

### Speaking tasks

Four of the eight computer-administered speaking tasks designed by De Jong *et al.* (in press) were used in this study. For each task, the instruction screens provided a photo picture of the communicative situation and one or several visual-verbal cues concerning the topic. Participants were invited to play the role of someone involved in the situation and to produce a monologue of not more than 2 min. In the study of De Jong *et al.* (in press), the speaking tasks were originally designed with contrasts on three dimensions: topic complexity (simple versus complex), formality (informal versus formal), and discourse genre (argumentative versus descriptive). Participants in the present study only performed the four complex speaking tasks:

- *Unemployment task* (informal, descriptive): Participant tells a friend about the development of unemployment among women and men over the past 10 years.



- *Transportation task* (informal, argumentative): Participant discusses the pros and cons of how to solve the problem of traffic congestions by using public transportation, bicycles, or automobiles.
- *Hospital task* (formal, descriptive): Participant works at the employment office of a hospital and tells a candidate for a nurse position what the main tasks in the vacant position are.
- *Car Park task* (formal, argumentative): participant is manager of a supermarket, addressing a neighbourhood meeting, arguing which one of three alternative plans for building a car park he/she prefers.

**Procedure.** Participants performed a familiarization task followed by the four test tasks presented in the same order for all participants. In each task, the communicative situation and additional information were presented in two consecutive screens. Then, a time bar appeared, filling blue in 30 s, indicating that the participants could prepare their monologue. Then another bar appeared, filling green in 120 s, forming the cue that the participants could start speaking and indicating the remaining speaking time that was left for speaking. Participants' responses were audio recorded and later transcribed following the Childes transcription conventions (MacWhinney 2000) as far as necessary.

**Scoring of communicative adequacy.** Being communicatively adequate or successful can be defined as achieving an outcome that is adequate in terms of the communicative goal one wants to achieve, with correct or appropriate propositional content. De Jong *et al.* (in press) constructed a rating scale with task-specific descriptors for each of the speaking tasks. This scale forms a task-specific operationalization of communicative efficiency scales used in language assessment, such as the scale 'provision of main ideas and supporting details' in the study of Iwashita *et al.* (2008). All descriptors were formulated in terms of functional speaking proficiency, distinguishing between differences in success in conveying the message. The rating scale comprised six levels, containing descriptors pertaining to (i) the amount and detail of information conveyed, relevant to the topic, setting (formal/informal) and discourse type (descriptive/argumentative) and (ii) the intelligibility of the response. Each level was subdivided in five intervals, thus creating a rating scale ranging from 1 to 30 (see Supplementary Appendix 1).

Three native speakers, university students of medicine, technical engineering and psychology, received payment to rate the audio recordings of the responses of all participants in all speaking tasks. After a short introductory training session, the judges received all responses on an audio disk for rating them at home. The order in which they rated the responses reflected the order in which the participants had done the tasks. Within each task, the 98 performances were ordered randomly for each rater.

**Scoring of linguistic features.** After all responses had been transcribed, the following variables were coded: total number of words (TW), total number of



T-units (TU), mean number of words per T-unit (WpTU), number of content words per T-unit (CWpTU) as a measure of lexical richness, number of hesitations per T-unit (HpTU), and grammatical errors per T-unit (GEpTU). In the TW count, expressions like *eh*, *ehm*, and *mm* were not included. However, incomplete words consisting of minimally one syllable were included in the count, as were exclamations like *hè*, *hé*, and *ha*, though the number of exclamations and incomplete words were very low. For the TU count, the transcripts were divided into T-units, using Hunt's (1965) definition of T-unit: a main clause plus any subordinate clauses. In the CWpTU variable we counted the total number of content words (nouns, verbs (excluding auxiliaries), adjectives and adverbs) and divided this total by the total number of T-units. The hesitations variable comprised the sum of false starts, repairs and repetitions divided by the number of T-units. Grammatical errors comprised the following violations to standard Dutch: incorrect verb, noun or adjective inflection, incorrect word order, incorrect omission of a word or constituent (determiner, conjunction, subject, verb, etc.), incorrect use of determiners or conjunctions, superfluous use of word or constituents, and incorrect use of *als* in a comparative construction. In spoken Dutch, the subject and finite verb are often omitted in a conversation. Such omissions were not counted as errors. Since many utterances in spoken language leave room for several interpretations, we counted as grammatical errors only obvious violations of grammar, ungrammatical even in spoken language. The GEpTU variable was computed by dividing the sum of grammatical errors by the number of T-units.

## RESULTS

### **Differences in lexical knowledge, lexical fluency, and lexical working memory capacity, as a function of age and level of EP**

The following mean (*M*) reaction times (RTs) and standard deviation (SD) values in milliseconds (ms) were obtained in the four lexical speed measures, for the sample as a whole: word association (*M* = 1353; SD = 333), visual lexical decision (*M* = 598; SD = 82), auditory decision (*M* = 806; SD = 81), and picture naming (*M* = 698; SD = 97). Accuracy in the two lexical knowledge tests was as follows: vocabulary (*M* = 68 per cent; SD = 15 per cent) and word association (*M* = 82 per cent; SD = 12 per cent). Accuracy in the two lexical working-memory tasks was as follows: visual word span (range = 3–8; *M* = 4.6; SD = 0.9) and auditory word span (range = 3–8; *M* = 4.8; SD = 0.9). To answer RQ 1, 3 × 2 ANOVAs were conducted on each of these variables with age (18–35, 36–50, and 51–76 years) and EP (low versus high) as the independent variables. Table 2 shows the results.

Age had a significant effect on all four speed measures. On average, senior subjects (Ss) responded 399, 105, 90, and 79 ms slower than the young Ss in the word association, visual lexical decision, auditory lexical decision, and picture naming tasks, respectively. Participants' age (i.e. their real age, not

Table 2: Probability levels of effects of age and EP (level of education and profession) on measures of lexical fluency, lexical knowledge, and lexical working memory

	Age	EP	Age × EP
Lexical fluency/speed			
RT word association	.001	NS	NS
RT visual lexical decision	.000	NS	NS
RT auditory lexical decision	.000	NS	NS
RT picture naming	.001	NS	NS
Lexical knowledge			
Vocabulary	.0022	.000	.030
Word association	NS	.048	NS
Lexical working memory			
Visual word span	.001	NS	.056
Auditory word span	.003	.034	NS

NS = not significant.

age group) was significantly associated with the four speed measures ( $r$  between .40 and .48, all coefficients  $p < .01$ ). Vocabulary knowledge differed by both age and EP, with younger Ss and low EP Ss obtaining lower scores than older Ss and high EP Ss, respectively. The Age × EP interaction was significant as well. As Figure 1 shows the interaction effect is mainly caused by the low scores of the low EP Ss in the youngest group.<sup>2</sup>

In the word association task, High EP Ss performed significantly better than low EP Ss (85 per cent versus 79 per cent). The effect of Age was significant in the two working-memory measures but the distance between the 18–35 and 51–76 year old Ss was not large (means of 5.0 versus 4.2 and of 5.2 versus 4.3 in visual and auditory word span, respectively). Finally, low EP Ss had a smaller auditory word span than high EP Ss (4.5 versus 5.0).

In summary, the answer to our first research question is that (i) Ss of older age performed more poorly than Ss of younger age on all measures of lexical fluency and lexical memory, but better on the vocabulary test, (ii) EP affected the two knowledge measures and one of the word span measures, while (iii) a significant Age × EP effect was obtained in vocabulary knowledge, in that the low EP Ss in the 18- to 36-year-old group scored much lower than the low EP Ss in the two older age groups.

A principal component analysis on the eight lexical measures was run to explore their structure. Three factors with eigenvalues larger than 1 emerged, together explaining 68 per cent of the variance, nicely fitting our original categorization into lexical fluency, lexical knowledge, and lexical working

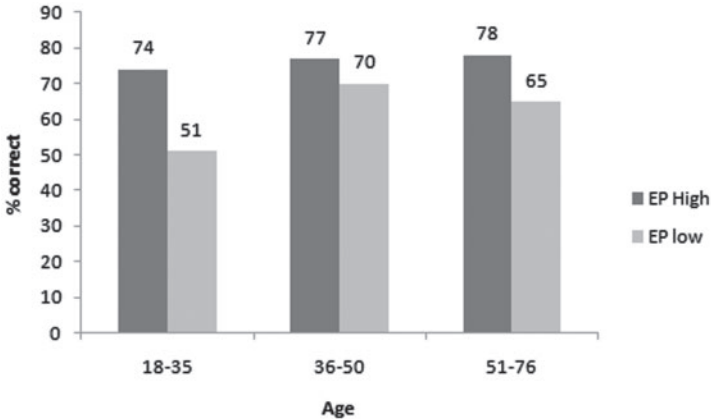


Figure 1: Vocabulary knowledge by age and EP

memory (Table 3). The best representatives of these three categories are word association and auditory lexical decision speed (equal contribution) for fluency, vocabulary for knowledge, and auditory word span for working memory. In some of the analyses to be reported below, auditory lexical decision, vocabulary, and auditory word span were used as independent variables, representing the three categories, respectively.

### Age and EP effects on communicative adequacy of performance in everyday speaking tasks

Because the quality of some of the recordings was too poor to allow coding and because a few participants did not perform all four speaking tasks, the  $n$ -values in the speaking measures range between 95 and 98. For 92 participants, judgments of all of the four tasks are available.

The four speaking tasks differed in difficulty with respect to the communicative adequacy (CA) with which they were performed, according to our rater panel. Inter-rater reliabilities (Cronbach's  $\alpha$ ) ranged between .77 and .86. On a 30-point CA scale with a score of 15 splitting the sufficient from the insufficient responses, the Unemployment task (informal, descriptive) turned out to be the most difficult one ( $M = 13.8$ ;  $SD = 5.5$ ), the Car Park task (formal, argumentative) the least difficult one ( $M = 17.7$ ;  $SD = 4.1$ ), while the Transportation task (informal, argumentative) ( $M = 15.1$ ;  $SD = 3.9$ ) and the Hospital task (formal, descriptive) ( $M = 15.8$ ;  $SD = 5.7$ ) did not differ substantially from one another in difficulty. A repeated-measures ANOVA on CA scores in the four tasks produced a main task effect:  $F(3,273) = 18.082$ ;  $p < .000$ ;  $\eta^2_p = .166$ .

The CA scores in each task were subjected to four  $3 \times 2$  ANOVAs, with Age and EP as independent variables. None of these ANOVAs produced an Age

Table 3: Factor structure of the measures of lexical fluency, lexical knowledge, and lexical memory

	Component		
	Lexical fluency	Lexical knowledge	Lexical memory
RT word association	<b>.841</b>	.160	.098
RT visual lexical decision	<b>.818</b>	-.143	-.095
RT auditory lexical decision	<b>.841</b>	.076	-.067
RT picture naming	<b>.601</b>	-.049	-.039
Percentage correct vocabulary test	-.118	<b>.889</b>	-.222
Percentage correct word association	.298	<b>.660</b>	.289
Mean visual word span	-.010	-.179	<b>.906</b>
Mean auditory word span	-.329	.251	<b>.606</b>

<sup>a</sup>Extraction method: principal component analysis; rotation method: Oblimin with Kaiser normalization.

effect or an Age  $\times$  EP interaction effect, while they all produced an EP effect ( $p < .01$ ). The age effect approached significance ( $p = .06$ ) in the Hospital and Car Park tasks.

### Linguistic aspects of speaking performances in relation to CA

Performance in the four speaking tasks was not only scored in terms of CA but also for several linguistic aspects: TW, TU, and, per T-unit, mean number of: words, content words, hesitations, and grammatical errors (respectively: WpTU, CWpTU, HpTU, and GEpTU). The four distributions of HpTU were skewed because several subjects (between 5 and 11) did not produce any hesitations. The highest HpTU score was 2.13. The distributions of GEpTU were also skewed, the number of subjects not producing any errors ranging between 2 and 7; the highest score being 2.33. We were surprised to find that a large majority of the participants made serious grammatical errors, such as violation of subject-verb agreement, in spontaneous speech.

Regression analyses were run, for each speaking task separately, with CA as the dependent variable and TW, TU, WpTU, CWpTU, HpTU, and GEpTU as the independent variables. Together, these variables explained 13, 59, 29, and 59 per cent of CA in the Unemployment, Transportation, Hospital, and Car Park tasks, respectively. Few variables produced contributions that were significant or approached significance: TW, WpTU, and CWpTU in the Transportation task ( $p = .000, .052, \text{ and } .007$ , respectively); TW, WpTU, CWpTU, and GEpTU in the

Hospital task ( $p = .067, .067, .019, \text{ and } .015$ , respectively); TW, CWpTU, and GEpTU in the Car Park task ( $p = .032, .012, \text{ and } .022$ , respectively); none of the predictors were significantly associated with CA in the unemployment task. We conclude that the CA ratings represent a genuine assessment of the information conveyed by the speaker, hardly confounded by hesitations or grammatical errors.

### Effects of age and education on linguistic aspects of speaking performances

The scores of TW, TU, WpTU, CWpTU, HpTU, and GEpTU, averaged over four tasks, were subjected to ANOVA with age and EP as independent variables. *Age effects* were obtained in none of the six analyses.<sup>3</sup> *EP effects* were obtained in three analyses: High EP Ss talked longer, producing more words ( $p = .003$ ;  $\eta^2_p = .092$ ), producing more T-units ( $p = .47$ ;  $\eta^2_p = .042$ ) and making fewer grammatical errors ( $p = .001$ ;  $\eta^2_p = .115$ ) than low EP Ss. However, no EP effects were found on words per T-unit, lexical richness (CWpTU) or hesitations per T-unit. Above, we already reported that the responses of high EP Ss were communicatively more successful than those of low EP Ss. We will return to this finding below.

The three judges who rated the CA of the responses in the speaking tasks had been instructed to focus on the quality and amount of information, as specified by the descriptions of the rating-scale values, while disregarding grammatical errors, speech hesitations or other response features if irrelevant to the CA of the responses. The CA ratings were indeed not confounded by the fluency or grammatical quality of the responses, as we demonstrated in the previous subsection. However, and not surprisingly, CA ratings were significantly associated with the number of words produced (coefficients of .26, .70, .46 and .71 in the Unemployment, Transportation, Hospital, and Car Park tasks, respectively,  $p < .00$  in all cases). With one exception, EP was significantly associated with both CA and number of words produced ( $r$  between .18 and .46), in all four tasks, meaning that high EP Ss tended to talk longer than low EP Ss. One might conclude then that high EP participants obtained higher CA scores because they talked longer. However, this simplistic conclusion is not warranted because the low (albeit often significant) coefficients between EP and CA and between EP and number of words clearly show that many long responses were produced by low EP participants and that many short responses were given a high CA rating.

### The association of CA of speaking with skills in lexical processing, lexical knowledge and lexical memory

While the regression analyses in which CA was regressed on TW, TU, WpTU, CWpTU, HpTU, and GEpTU provide, as it were, an internal picture of speaking proficiency, featuring linguistic predictors of the performances of which CA

was assessed, the regression analyses to be reported now aim at predicting CA in speaking with linguistic skills measured independently. We regressed CA, in each of the four speaking tasks, on (i) auditory lexical decision speed, representing lexical fluency, (ii) vocabulary knowledge, representing lexical knowledge, and (iii) auditory word span, representing lexical working memory. The variance of CA explained by these three skills was 16, 12, 22, and 25 per cent in the Unemployment, Transportation, Hospital, and Car Park tasks, respectively. Significant contributions were produced by lexical decision ( $t = -2.2$ ;  $p = .029$ ) and vocabulary ( $t = 3.0$ ;  $p = .003$ ) in the Unemployment task, by word span ( $t = 2.8$ ;  $p = .006$ ) in the Transportation task, by lexical decision ( $t = -2.0$ ;  $p = .04$ ), vocabulary ( $t = 3.6$ ;  $p = .000$ ), and word span ( $t = 2.0$ ;  $p = .04$ ) in the Hospital task, and by lexical decision ( $t = -3.6$ ;  $p = .000$ ) and vocabulary ( $t = 3.5$ ;  $p = .001$ ) in the Car Park task.

CA scores in the four speaking tasks were significantly associated with one another ( $p < .00$ ), with  $r$  ranging between .46 and .55. A principal component analyses on the four CA scores produced a single component, explaining 58 per cent of the variance; the four task loadings ranged between .70 and .83. The factor scores were saved, constituting combined CA scores across the four speaking tasks. In order to examine to what extent CA in speaking was associated with the linguistic skills measured separately, the combined CA scores were regressed, in a step-wise manner, on (i) vocabulary knowledge, representing lexical knowledge, (ii) auditory lexical decision speed, representing lexical fluency, and (iii) auditory word span, representing lexical working memory. All three predictors contributed significantly. Vocabulary knowledge (Model 1) explained 15 per cent of the variance in the combined CA scores. Addition of auditory lexical decision speed (Model 2) increased the explained variance to 27 per cent, while addition of auditory word span (Model 3) produced a further increase to 31 per cent. Thus, lexical fluency, lexical knowledge, and lexical memory skills were shown to be significantly associated with the adequacy of the information conveyed in the four speaking tasks, together explaining 31 per cent of the variance.

### **Educational level and vocabulary knowledge as predictors of CA of speaking**

The combined CA score was regressed, in an additional analysis, not only on EP but also on the three linguistic skills. (Age was not included in this analysis because, as we reported above, no age effect on CA was found in any of the four tasks.) EP explained 26 per cent of the variance in CA ( $t = 5.589$ ;  $p = .000$ ); through the addition of vocabulary this percentage was raised to 29 per cent, and through the addition of lexical decision and then word span the percentage of explained variance increased to 36 and 37 per cent, respectively. In the final analysis, all predictors except word span contributed significantly (EP:  $t = 2.984$  with  $p = .004$ ; vocabulary:  $t = 2.394$  with  $p = .019$ ; auditory lexical decision:  $t = -2.735$  with  $p = .008$ ; auditory word span:  $t = 1.454$  with  $p = .15$ ).



Thus, although vocabulary knowledge was associated (not surprisingly) to EP (biserial  $r = -.48$ ;  $p = .000$ ), vocabulary knowledge was not entirely implicated in EP because it explained 3 per cent of the variance in CA on top of EP, while even speed of lexical processing (lexical decision) contributed uniquely for another 7 per cent.

## DISCUSSION

Following Escudero and Sharwood Smith's (2001) call for further research on the notion of linguistic ability in native speakers, our study assessed, in a sample of 98 adult NSs of Dutch, how their lexical skills and their speaking proficiency varied as a function of their age and level of EP. Participants were categorized in terms of their age (18–35, 36–50, and 51–76 years old) and the level of their EP (low versus high). In this section, we first summarize the main findings and then discuss the findings, rounding off with a conclusion.

### Summary of the main findings

Lexical skills were assessed with seven tasks producing four measures of lexical fluency (reaction times), two measures of lexical knowledge, and two measures of lexical memory. Lexical fluency was affected by Age but not by EP; older participants responded more slowly than younger participants. Vocabulary knowledge increased with age and was higher in the high EP group than the low EP group; the Age  $\times$  EP interaction was significant as well. The difference in vocabulary knowledge between low and high EP subjects was relatively large in the 18–35 year group but smaller in the two older age groups. Lexical memory in the aural mode was affected by both Age and EP, favouring the younger and more highly educated participants; in the visual mode, significant effects were obtained for age and the Age  $\times$  EP interaction (RQ1).

Participants with higher EP were more successful in conveying their messages in the four speaking tasks than participants with lower EP but age was not found to be associated with communicative success (RQ2).

Participants with higher EP talked longer, producing more words and T-units and making fewer grammatical errors than Low EP participants. However, EP was not found to be associated with words per T-unit, lexical richness, or hesitations per T-unit. No age differences were found in response length, lexical richness, hesitations, or grammatical errors (RQ3).

A regression analysis, in which communicative success in speaking (a combined score across four speaking tasks) was regressed on three lexical skills, showed that 15 per cent of the variance in communicative success was explained by vocabulary knowledge, an additional 12 per cent by auditory lexical decision (a speed variable), and, on top of that, 4 per cent by auditory word span. Thus, lexical knowledge, the speed with which lexical knowledge can be processed, and the capacity for keeping lexical information in a memory buffer

for a short period, all contributed significantly to the information conveyed in the speaking tasks (RQ4).

A similar step-wise regression analysis, this time with EP as the first predictor, was conducted to see to what extent the lexical skills were implicated in the EP variable in explaining CA in speaking. This analysis showed that communicative success in speaking was contingent on EP (26 per cent), vocabulary knowledge (additional 3 per cent), speed in auditory lexical decision (additional 7 per cent), and auditory word span (additional 1 per cent). Thus, lexical knowledge, lexical fluency, and lexical memory were significantly associated with the quality of information provided in the speaking tasks even beyond the association with level of EP (RQ5).

## Discussion

The results of our study largely support the findings of the psycholinguistic studies reviewed at the beginning of this article with respect to the effect of age. Older participants responded more slowly in the four lexical speed tasks, more poorly in the two word span tasks, but better in the vocabulary knowledge task (Table 2). However, older and younger participants performed equally well in the speaking tasks. The only task in which main effects of both age and EP as well as an Age  $\times$  EP interaction were found, was the vocabulary knowledge task. The age and Age  $\times$  EP effects were caused by a relatively low performance in the low EP group of 18- to 35-year-olds. Although it seems plausible that adults increase their vocabularies across the life span, we have no explanation why this should be more so for low EP people than for high EP people. Clearly, more research is needed to examine the robustness of this finding and to offer an adequate explanation.

The added value of our study to the empirical literature reviewed at the beginning of this article, resides, we would like to argue, in the fact that we tested participants on a variety of lexical subskills (knowledge, speed of processing, and span of processing) as well as on their ability to produce meaningful speech, representing descriptive and argumentative discourse in informal and formal communicative situations. Through this combination, we were able to demonstrate how the communicative quality of speech produced in such settings is associated with lexical skills. This association was obtained not only in the case of lexical knowledge, a finding that might not surprise the research community, but also with respect to the speed with which lexical knowledge can be processed and the memory capacity for lexical information, beyond the level of education and vocabulary size. The findings suggest that although older people (not suffering from severe mental impairment) appear to process lexical information more slowly than younger people and although older people appear to have a smaller lexical memory span than younger people, they are not less successful in getting their message across in speaking. More research is needed, however, to support this generalization.

An unexpected finding of this study is that most participants, at both levels of EP, produced, in their responses in the speaking tasks, grammatical errors constituting obvious violations to the grammar of Dutch, such as violations of noun gender and violations of subject–verb agreement (number agreement).

Another unanticipated finding concerns the ratings of the content (CA ratings) of the responses in the speaking tasks. Though we did find education-related differences with respect to the quantity and quality of speaking performances, which is in line with earlier research on educational differences in picture naming (e.g. Beland *et al.* 1993), we observed that many of our participants, and not only those from the low EP group, failed to supply all required information in the descriptive tasks (although all information was given in the stimulus materials) or failed to provide well developed reasons in the argumentative tasks. We had expected the Car Park task to be the most difficult one and the Unemployment task the easiest one because of their contrast in both formality and discourse type. It turned out, however, that, in the Unemployment task, the graph representing the unemployment figures for men and women over time was difficult to grasp for people with low EP, while the Car Park task was rather easy because verbal cues for the three alternative plans were already given in the stimulus information and subjects only had to repeat them. To find arguments for one of the plans turned out to be not so difficult, not even for subjects with low EP. However, a panel of six experts in the area of L2 proficiency assessment, had rated the Unemployment task at the B1 level of the Common European Framework of Reference for Languages (Council of Europe 2001), one level lower than the other three tasks, which had been rated as B2 (see Hulstijn *et al.* under review). What these facts appear to suggest, is that ‘we’ language experts tend to underestimate the difficulty of test tasks and that, in L2 assessment, test tasks should be piloted on native speakers before they be labelled with a value on the CEFR scale for non-native speakers.

## CONCLUDING REMARKS

Our study aimed at improving our understanding of language ability in adult native speakers, as recommended by Escudero and Sharwood Smith (2001). Our findings, however, will not make it easier to define what Escudero and Sharwood Smith call the prototypical native speaker. Earlier studies had already shown that native speakers’ intuitions concerning the grammaticality of certain sentences (Chipere 2001) and native speakers’ comprehension of sentences (Dabrowska 1997) may differ as a function of their level of education. Our findings point to substantial differences among native speakers both in linguistic subskills and in speaking proficiency, suggesting that it is impossible to define the prototypical native speaker in terms of language ability. We presume that such differences reflect the level and amount of verbal activities in people’s daily lives, of which level of education and level of profession may only form an imperfect index. Future studies should investigate whether

speaking proficiency can be better predicted by indexes sensitive to the amount and type of verbal activities (in the oral and literate domain) than by our gross indexes of level of education and level of profession.

Our conclusion that the language ability of the prototypical native speaker cannot be simply defined may be relevant to the age question in L2 acquisition, which is commonly phrased as follows: Is it possible for L2 learners to reach native or near-native levels of proficiency when starting to learn the L2 after a certain (critical) age? In empirical studies of this issue (see reviews by Hyltenstam and Abrahamsson 2003; Birdsong 2006), L2 learners are typically tested on some measures of L2 proficiency, comparing their performance with that of NSs. But who are the NSs in such studies in terms of age and level of education?

As Hulstijn (in press) observed 'since Chomsky (1965: 4) claimed that all adult native speakers share the same grammatical competence [...], most researchers have simply taken the proposition for granted, neglecting the obligation of finding out to what extent it can be empirically upheld'. We, therefore, recommend that L2 learners in language assessment, and in studies investigating the critical-period issue be compared with NSs at the same educational or professional level as their own, and perhaps with NSs of the same age.

## SUPPLEMENTARY DATA

Supplementary material is available at *Applied Linguistics* online.

## ACKNOWLEDGEMENTS

At the time of the study, the first author was a Masters student and a junior researcher at the University of Amsterdam, conducting the work reported here under the supervision of the second author. We would like to thank the Board of Directors of *Language Learning* for awarding the 2009 *Language Learning* Research Grant to the second author, allowing the first author to complete this study, after she had obtained her Masters degree in the linguistics programme.

## NOTES

- 1 Education and profession were combined in a way not unlike the variable socio-economic status (SES) widely used in the socio-economic literature. However, whereas SES includes income, our EP variable deliberately excludes income.
- 2 A univariate ANOVA on vocabulary knowledge, with EP as independent factor and Age (subjects' real age) as covariate, produced a significant EP effect ( $df=1.95$ ;  $p=.000$ ;  $\eta^2_p=.273$ ) and a significant Age effect ( $df=1.95$ ;  $p=.004$ ;  $\eta^2_p=.085$ ).
- 3 Subjects' age (i.e. their real age, not age group) did not correlate significantly with TW, TU, WpTU, CWpTU, HpTU, or GEpTU in any of the four speaking tasks, with three exceptions:  $r=.27$  and  $r=.23$  between Age and GEpTU in the Hospital and Car Park tasks, respectively (both  $p=.001$ ), and  $r=-.25$  between Age and CWpTU in the Car Park task ( $p=.013$ ).

## REFERENCES

- Anderson, J.** 1980. *Cognitive Psychology and its Implications*. Freeman.
- Beland, R., A. R. Lecours, F. Giroux, and M. Bois.** 1993. 'The MT-86  $\beta$  Aphasia Battery: A subset of normative data in relation to age and level of school education (Part II),' *Aphasiology* 7/4: 359–82.
- Birren J. E. and K. W. Schaie** (eds). 2006. *Handbook of the Psychology of Aging*, 6th edn. Academic Press.
- Birdsong, D.** 2006. 'Age and second language acquisition and processing: A selective overview,' *Language Learning*, 56(Suppl. 1): 9–49.
- Bosma, H., M. P. J. Van Boxtel, R. W. H. M. Ponds, P. J. Houx, and J. Jolles.** 2003. 'Education and age-related cognitive decline; the contribution of mental workload,' *Educational Gerontology* 29: 165–73.
- Burke, D. M., D. G. MacKay, and L. E. James.** 2000. 'Theoretical approaches to language and aging' in T. Perfect and E. Maylor (eds): *Models of Cognitive Aging*. Oxford University Press, pp. 204–37.
- Centraal Bureau voor de Statistiek.** 2001. Standaard Beroepenclassificatie. Voorburg/Heerlen.
- Chipere, N.** 2001. 'Variations in native speaker competence: Implications for first-language teaching,' *Language Awareness* 10: 107–24.
- Chomsky, N.** 1965. *Aspects of the Theory of Syntax*. MIT Press.
- Conway, A. R. A., C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse.** 2007. in A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, and J. N. Towse (eds): *Variation in Working Memory*. Oxford University Press, pp. 3–17.
- Council of Europe.** 2001. Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Cambridge University Press/Council of Europe.
- Craik F. I. M. and T. A. Salthouse** (eds). 2008. *The Handbook of Aging and Cognition*, 3rd edn. Psychology Press.
- Dabrowska, E.** 1997. 'The LAD goes to school: a cautionary tale for nativists,' *Linguistics* 35: 735–66.
- Davies, A.** 2003. *The Native Speaker: Myth and Reality*. Multilingual Matters.
- De Bot K. and R. W. Schrauf** (eds). 2009. *Language Development over the Lifespan*. Routledge.
- De Groot, A. M. B.** 2011. *Language and Cognition in Bilinguals and Multilinguals*. Psychology Press.
- De Jong, N. H., M. P. Steinel, A. F. Florijn, R. Schoonen, and J. H. Hulstijn.** 2011. 'Facets of speaking proficiency,' *Studies in Second Language Acquisition* (in press).
- De Kleijn, P. and E. Nieuwborg.** 1996. *Basiswoordenboek Nederlands*. Wolters.
- Escudero, P. and M. Sharwood Smith.** 2001. 'Reinventing the Native Speaker: Or "What You Never Wanted to Know about the Native Speaker So Never Dared to Ask"', *EUROSLA Yearbook* 1: 275–86.
- Hawkins, K. A. and S. Bender.** 2002. 'Norms and relationship of the Boston Naming Test performance to vocabulary and education: A review,' *Aphasiology* 16: 1143–53.
- Hulstijn, J. H.** 2011. 'Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second-language assessment,' *Language Assessment Quarterly* (in press).
- Hunt, K. W.** 1965. *Grammatical Structures Written at Three Grade Levels*. National Council of teachers of English.
- Hyltenstam, K. and N. Abrahamsson.** 2003. 'Maturational constraints in SLA' in C. J. Doughty and M. H. Long (eds): *The Handbook of Second Language Acquisition*. Blackwell, pp. 539–99.
- Iwashita, N., A. Brown, T. McNamara, and S. O'Hagan.** 2008. 'Assessed levels of second language speaking proficiency: How distinct?,' *Applied Linguistics* 29: 24–49.
- Le Dorze, G. and C. Bedard.** 1998. 'Effects of age and education on the lexico-semantic content of connected speech in adults,' *Journal of Communication Disorders* 31: 53–71.
- Mackenzie, C.** 2000. 'Adult spoken discourse: The influences of age and education,' *International Journal of Language and Communication Disorders* 35: 269–85.
- Mackenzie, C., M. Brady, J. Norrie, and N. Poedjianto.** 2007. 'Picture description in neurologically normal adults: Concepts and topic coherence,' *Aphasiology* 21/3: 340–54.
- MacWhinney, B.** 2000. *The CHILDES Project: Tools for Analyzing Talk*, 3rd edn. Lawrence Erlbaum Associates.

- Meijer, A., H. M. De Groot, P. J. Van Boxtel, W. M. Van Gerven, and J. Jolles.** 2008. 'Are age differences in verbal learning related to interstimulus interval and education?,' *Experimental Aging Research* 34: 323–39.
- Paradis, M.** 2004. *A Neurolinguistic Theory of Bilingualism*. Benjamins.
- Patricacou, A., E. Psallida, T. Pring, and L. Dipper.** 2007. 'The Boston Naming Test in Greek: Normative data and the effects of age and education on naming,' *Aphasiology* 21/12: 1157–70.
- Poelmans, P.** 2003. 'Developing Second-language Listening Comprehension,' Doctoral thesis, University of Amsterdam.
- Salthouse, T. A.** 2009. 'When does age-related cognitive decline begin?,' *Neurobiology of aging* 30: 507–14.
- Schneider, A., M. Daneman, and M. K. Pichora-Fuller.** 2002a. 'Listening in aging adults: From discourse comprehension to psychoacoustics,' *Canadian Journal of Experimental Psychology* 56/3: 139–52.
- Schneider, W., A. Eschman, and A. Zuccolotto.** 2002b. *E-prime User's Guide and E-prime Reference Guide*. Version 1.0. Psychology Software Tools, Inc.
- Schrauf, R. W.** 2008. 'Bilingualism and aging' in J. Altarriba and R. R. Heredia (eds): *An Introduction to Bilingualism: Principles and Processes*. Erlbaum, pp. 105–27.
- Schrooten, W. and A. Vermeer.** 1994. *Woorden in het Basisonderwijs 15.000 Woorden Aangeboden aan Leerlingen*. Tilburg University Press.
- Tun, P. A. and A. Wingfield.** 1999. 'Language and communication: Fundamentals of speech communication and language processing in old age' in A. D. Fisk and W. A. Rogers (eds): *Handbook of Human Factors and the Older Adult*. Academic Press, pp. 125–49.
- Ullman, M. T.** 2004. 'Contributions of memory circuits to language: The declarative/procedural model,' *Cognition* 92: 231–270.
- Van der Elst, W., M. P. Van Boxtel, G. J. Van Breukelen, and J. Jolles.** 2005. 'Rey's verbal learning test: Normative data for 1855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation,' *Journal of the International Neuropsychological Society* 11: 290–302.
- Waters, G. and D. Caplan.** 2005. 'The relationship between age, processing speed, working memory capacity, and language comprehension,' *Memory* 13/3–4: 403–13.