Interpreters – experts in careful listening and efficient encoding? Findings of a prose recall test

$\circledS$SAGE

Sinikka Hiltunen
University of Helsinki, Finland

Gun-Viol Vik
University of Vaasa, Finland

Abstract

Research Questions: The purpose of the present experiment was to study interpreters' recall of spoken prose.
Design: The prose recall of simultaneous and consecutive interpreters was compared to that of foreign language teachers and non-linguistic experts. The professional experience of participants (21–24 participants in each group) was 10 years as a aminimum. The auditory presentation of the prose passage to be recalled, divided into eleven speech sequences, resembled the working conditions of interpreters.
Data: Transcripted prose recall recordings were analysed quantitatively through an idea unit measure and qualitatively through meaning-based expressions.
Findings: The foreign language expert groups outperformed the non-linguistic experts in both quantitative and qualitative measures. Additionally, compared to foreign language teachers, interpreters indicated a better recall of time expressions and topic sentences, as well as of complicated emotional and causal expressions. The explanation for these findings could indicate expertise-dependent tendencies: possibly a continuous practising of careful listening and the demand for a quick comprehension of the source text under the extreme time pressure of interpreters' work leads to better results in prose recall. However, the findings can only be generalized to a limited extent because the prose passage used contained only one or two expressions of each type studied in the qualitative analysis.
Originality: The study differs from previous studies in that the memory of interpreters, and especially of consecutive interpreters, was studied for the first time with a prose recall measure.
Significance: The prose recall test revealed that the abilities of careful listening and effective comprehension of coherence and causality seem to play a significant role in explaining memory functions of simultaneous and consecutive interpreters compared to those of foreign language teachers and non-linguistic experts.

Keywords

Careful listening, consecutive interpreter, foreign language teacher, prose recall, simultaneous interpreter

Corresponding author:
Sinikka Hiltunen, University of Helsinki, P.O. Box 9, Helsinki, FI-00014, Finland.
Email: sinikka.hiltunen@muistikuisti.net

Introduction

The main objective of the present study was to assess the memory of expert interpreters, both simultaneous and consecutive, through prose recall measures. The study falls into the fields of cognitive science and particularly of research of expert memory. The prose recall experiment formed the second part in a series of experiments on the memory and executive functions of simultaneous and consecutive interpreters. The results of the first part have been reported by Hiltunen, Vik, Pääkkonen and Krause (2016). The experiments were designed for expert participants with a minimum of ten years of professional experience, as defined by Ericsson and Kintsch (1995; for interpreters, see Obler, 2012). To the best of our knowledge, our series of experiments represents the first time that the memory and executive functions of consecutive interpreters have been studied.

By definition, expert performance is 'consistently superior performance on a specified set of representative tasks for a domain' (Ericsson & Smith, 1991). Achieving an expert level of performance requires continuous deliberate practice. This includes, among others, the acquisition, application and organization of knowledge into specialized patterns for future use, as well as the continuous modification of patterns on the basis of feed-back (Ericsson & Lehmann, 1996; Ericsson, Brampe & Tesch-Römer, 1993). As a result, experts can outperform novices in memory tests when the materials presented are organized in a meaningful way, as has been indicated by research on several groups of experts, such as taxi drivers (Kalakoski & Saariluoma, 2001), architects and teachers of native language and literature (Cavallini, Cornoldi, & Vecchi, 2009), and sign language interpreters (Macnamara, Moore, Kegl, & Conway, 2011), to name a few. Similar results have also been demonstrated by research on the memory of simultaneous interpreters (for a review, see, e.g. García, 2014; Signorelli, Haarmann, & Obler, 2012).

Deliberate practice leading to better memory results in experiments is an effortful activity motivated by the goal of continuously improving one's performance (Ericsson & Charness, 1994). In this way, expertise in any field can be refined on a proficiency scale from novice to journeyman, and further to expert and master, the last two characterizing the level of the best professionals in the field and even exceeding it (Chi, 2003, adapted from Hoffman, 1998).

However, for some groups of experts at least, deliberate practice does not seem to cover the main skills in the domain. In her dissertation, Tiselius (2013) showed that the simultaneous interpreters interviewed by her more often deliberately practised interpreting sub-skills by improving their language skills, adding to their general knowledge, and actively learning from their peers by listening to their performance at work. It can be presumed that not only these activities, but also the very fact of working as a professional interpreter should improve the memory skills of interpreters.

As to research methods, Ericsson and Williams (2007) have recommended that research on experts' memory should include at least the following: first, identifying significantly superior performance and then designing conditions for eliciting this performance in a controlled laboratory situation. Consequently, the present prose-recall experiment was designed to resemble, as far as possible, the working conditions of interpreters (see Materials and methods for details).

According to the definition by Pöchhacker (2004, 11; emphasis in the original) 'Interpreting is a form of Translation in which a **first and final rendition in another language** is produced on the basis of a **one-time presentation** of an utterance in a source language.' Implicitly, this definition emphasizes one of the many expert skills that have to be thoroughly learned and practised by interpreters: the ability to listen carefully. For simultaneous interpreters, this means dividing their attention between several cognitive functions under an extreme time pressure: careful listening, transcoding, speech production, and monitoring the equivalence between the source language and target language texts. Consecutive interpreters do not experience the same time pressure, but they do have to concentrate on careful listening, encoding, and maintaining the message (possibly using notes), and rendering it in the target language when the speaker pauses. In some cases, it is possible

for consecutive interpreters to ask the speaker for clarification if something has been missed or not completely understood while listening. This, however, should not happen too often. Simultaneous interpreters do not have this possibility, which makes careful listening even more important for them. Nonetheless, for both groups of interpreters, careful listening is essential in order to effectively encode the message and reproduce it in the target language.

In measuring the capacity of the working memory, memory research has traditionally used reading or listening span tests, among others. In a listening span test, the participant has to listen to unrelated sentences and after the test, to recall the last word of each sentence in serial order. According to theory, the test measures different components of working memory (for components, see, e.g. Baddeley, 1986, 1990): comprehension of sentence content for later question-answering, classified as processing, and encoding and later recall of last words, classified as storage or maintenance (Daneman & Carpenter, 1980). The listening span has been found to correlate with the reading span[1] (Baddeley, Logie, & Nimmo-Smith, 1985; Daneman & Merikle, 1996) and with listening comprehension (Daneman & Carpenter, 1980).

The working memory of simultaneous interpreters has been tested both with listening span and reading span tests. In the following review, although, only studies on professional interpreters with enough experience (averaging 10 years or more, see, e.g. Signorelli et al., 2012) and using native-language materials are included (as is the case in the present study). For a more comprehensive review, see García (2014).

The reading span test by Christoffels, de Groot, and Kroll (2006) suggested a significant difference between expert interpreters and foreign language teachers, but the difference between professional interpreters and bilingual university students (novices) did not reach significance. Furthermore, the studies by Yudes, Macize, and Bajo (2011) indicated a higher reading span for the simultaneous interpreters than for novices (monolingual or bilingual participants). In contrast, in the listening span test by Köpke and Nespoulous (2006), beginning interpreters outperformed both experienced interpreters and two control groups (bilingual adults and students), but the difference in recall was significant only between the beginning interpreters and both control groups.

Thus, the reading span measures seem to indicate an advantage for simultaneous interpreters, whereas the listening span measure does not. This is somewhat surprising, considering that interpreters, in particular, should be more accustomed to listening than other (bilingual) groups. For additional aspects of this dilemma, see General discussion.

With this in mind, instead of listening or reading span measures, we preferred a modified prose recall test with an auditory presentation of passages. Previously, prose recall with visually presented materials has been used as part of memory research on experts by Shimamura, Berry, Mangels, Rusting, and Jurica (1995). They compared professors of different age groups to standard old and standard young participants in the recall of prose passages involving popular scientific and historical-anthropological information, targeted to non-experts. The results demonstrated that professors of all ages outperformed both standard young and old participants.

The effect of presentation modality on the accuracy of prose recall seems to be twofold. On the one hand, listening comprehension is an important factor predicting reading comprehension (Macaruso & Shankweiler, 2010). On the other hand, listening seems to suffer from some disadvantages compared to reading. For instance, prose recall has proved to be weaker with an auditory presentation than with a visual one for students (Cornoldi & De Beni, 1991), and across both young and old age groups (Corgiat, Templer, & Newell, 1989). Still, listening comprehension is considered by many authors to involve the same cognitive processes as reading comprehension (see, e.g. Jackson and McClelland, 1979; Macaruso & Shankweiler, 2010; Palmer, MacLeod, Hunt, & Davidson, 1985; for simultaneous interpreters, Dillinger, 1994), and the correlation in comprehension between written and auditory stories seems to be high: $r = 0.92$ (Gernsbacher, Varner, & Faust, 1990). Consequently, most of the findings in studies on the reading comprehension of prose can probably be applied even to listening comprehension.

Several researchers of reading comprehension emphasize the maintenance of coherence through inferring causal relations at local and global levels (Fletcher and Bloom, 1988; Fletcher et al., 1990; Graesser, Singer, & Trabasso, 1994; Kintsch, 1972; O'Brien & Albrecht, 1991). A coherent text allows the reader to make inferences and access antecedents (O'Brien & Albrecht, 1991). Moreover, the number of causal connections and grasping the causal chain predict immediate and delayed recall[2] (Trabasso & van den Broek, 1985). Since interpreters are specialized in listening, translating and producing spoken discourse, their ability to understand and reproduce causal relations and maintain coherence should be especially well developed. Accordingly, Jiang, Han, Jiang, and Feng (2014), for instance, have also emphasized coherence and topic continuity in their modern Triadic Discourse Interpreting Model (TRIM).

Still, according to Dillinger (1994), comprehension as demonstrated by simultaneous interpreters is not a specialized ability, but an ability that seems to accompany bilingualism and to be limited by the same parameters that limit comprehension in general: the nature of the text and the reader's prior knowledge (cf. Shimamura et al., 1995). Despite this, it could be presumed that the experience and practice in listening acquired by interpreters could lead to advanced abilities in prose recall. However, as this was the first time that the memory of interpreters was studied with a prose prose recall test, no advance hypotheses were formulated.

Material and methods

Participants

Ninety-one participants volunteered. The participants consisted of four groups of experts: three groups of foreign language experts and one group of non-linguistic experts. The three language expert groups were: simultaneous ($n = 21$, all female) and consecutive ($n = 22$; 20 female, 2 male) interpreters and foreign language teachers ($n = 24$; 23 female, 1 male). The group of non-linguistic experts consisted of 24 participants (19 female, 5 male). Owing to technical difficulties, the results of one consecutive interpreter (male) had to be excluded from the analyses.

Two criteria for expertise were used: level of education (B.A. or equivalent as a minimum), and at least 10 years of experience in one's field of expertise (Ericsson & Lehmann, 1996; Obler, 2012); for details, see Table 1. There were no significant differences between the four groups as regards age ($F(3,86) = 81.358$, $p= 0.232$) or professional experience ($F(3,86) = 0.460$, $p= 0.711$). All participants spoke Finnish as their native language or were early bilinguals (Finnish learned before the age of 5). No additional data of other languages were gathered because the materials presented were in the native Finnish. For occupational data, see Table 1.

Materials and procedure

The prose recall test took place in a counterbalanced order with the free recall test between the participants (see Table 2; also, for information on procedure details, see Hiltunen et al., 2016). For the prose recall test an extract from an on-line article 'Forced Loneliness', ('Pakotettu yksinäisyys', Raitasalo, 2006) was chosen, see Appendix 1. The topic is frequently dealt with in both popular mass media and more specialized publications, and the article was published in a web magazine targeted to a very wide audience. More specifically, it did not contain words, concepts or information which would have been difficult to understand by a mother-tongue listener. The participants were not informed in advance of the topic, but were told that the passage was a 'short description of the results of a research study'.

The length of the passage was 186 words. The text was recorded and edited with Cool Edit 2000 (by Syntrillium Software Corporation), spoken by a female voice. The text was divided into eleven speech sequences with enough time between the sequences for spoken recall. The sequence length

Table 1. Participants. Mean age, professional experience and other background information.

Group	n	Mean age, years (SD)	Education level	(Field of) occupation	Mean professional experience, years (SD)
Simultaneous interpreters	21	47.6 (7.6)	B.A. or equiv. (2) M.A. (19)	conference interpr. (16) court interpreter (11) interpr. in business negotiations (10) community interpr. (9) translator (10) teacher (2)	16.1 (6.6)
Consecutive interpreters	22	49.6 (7.5)	B.A. or equiv. (9) M.A. (13)	community interpreter (16) court interpreter (13) interpr. in business negotiations (13) conference interpreter (2) translator (6) teacher (9) personal assistant (2)	17.4 (8.5)
Foreign language teachers	24	47.1 (7.0)	B.A. or equiv. (1) M.A. (23)	foreign language teacher (24) head teacher (2) researcher (1)	18.0 (7.1)
Non-linguistic experts	24	44.9 (7.8)	B.A. or equiv.(19) M.A. (5)	administration and clerical work (9) manufacturing industry (5) health care (3) teaching (other than languages) (3) commerce (2) culture (1)	19.4 (9.8)

Table 2. Experimental procedure: testing order and filling in forms for background information (for other tests in the series and their procedure details, see Hiltunen et al., 2016).

Practise free recall test with 10 abstract words
Filling in an inquiry on personal beliefs regarding memory and attention capabilities
Experimental free recall test with abstract words
Filling in a form on memory strategies used in the free recall test
Prose recall test (in a counter-balanced order with the free recall test between the participants)[a]
Filling in a form on memory strategies used in the prose recall test
Filling in a questionnaire on background information (if not submitted beforehand via e-mail)
Filling in an inquiry on personal beliefs regarding memory and attention capabilities

[a]Only the prose recall test is reported in this paper.

varied from 8 to 23 words (5.0 to 16.2 s). On average, assuming natural pauses between the speech sequences, the text was spoken at 89 words/min.[3]

  The participants listened to the speech sequences through headphones and were instructed to recall each sequence aloud to the best of their ability. The importance of maintaining the message was emphasized (as is usual in interpreting), but the participant was also told that there would be

extra points for details correctly recalled. The stimuli were presented from a CD player or MP3 player, and the spoken answers were recorded with a laptop computer, using Cool Edit 2000.

Analyses

The recordings were transcribed into .doc format and the statistical analyses were carried out with SPSS for Windows 21.0. For the prose text, the idea unit as defined by Mills, Diehl, Birkmire, and Mou (1993) was chosen as the appropriate unit of analysis. The idea unit refers to 'a sentence or part of a sentence that expresses a complete idea which contains an actual or implicit verb and is usually a phrase-size unit' (Mills et al., 1993, p. 289). This analysis unit was used because it is mainly the sense of the expression that should be conveyed in interpreting (Seleskovitch, 1976), rather than words (or propositions, which are traditionally used in analysing prose recall). Idea units have been used earlier in several other prose recall and similar studies (e.g. Blunt & Karpicke, 2014; Corgiat et al., 1989; Glover, 1981; Lustig, May, & Hasher, 2001; Marschark, De Beni, Polazzo, & Cornoldi, 1993; McDaniel, 1984; McDaniel & Kerwin, 1987; Miangah, 2014; Sisco, Marsiske, Gross, & Rebok, 2013).

The texts were parsed into idea units by the main author (experimenter) together with an independent judge who herself has worked as simultaneous and consecutive interpreter as well as in interpreter education. Two points were given for each correctly recalled idea unit (see the definition above), with a maximum score of 62 points for the entire text (see Appendix 1). No points were removed for errors or incorrectly recalled idea units.

The results of 11 randomly selected participants (5 interpreters and 6 teachers) were analysed by the independent judge mentioned earlier, in addition to the experimenter. The reliability (Cronbach's alpha) of the analyses between the two judges was high: 0.93. The results reported below are based on the analysis made by the main author, although in many cases after consultation with the other judge.

Results, quantitative (statistical) analysis

The probability of prose recall by group is shown in Table 3. The mean recall of prose in idea units out of the maximum possible was 0.84 (SD 0.09) for simultaneous interpreters, 0.81 (SD 0.08) for consecutive interpreters, 0.77 (SD 0.07) for foreign language teachers, and 0.68 (SD 0.12) for non-linguistic experts.

To analyse the statistical significance of between-group differences, ANOVA (analysis of variance) was used. The criterion for statistical significance was set at $p = 0.05$ for all subsequent analyses. The results revealed a significant between-group effect in the probability of prose recall, $F(3,86) = 13.172$, $p < 0.001$, $\eta_p^2 = 0.315$. The Bonferroni corrected contrast analysis demonstrated that the significant difference was due to all three language expert groups outperforming the non-linguistic experts, with significance values of $p < 0.004$ for all groups.

To reveal the possible effect of speech sequence length, the difference between the short (11 words or less) and long (17 words or more; see Table 3) sequences was analysed with a 2-within (Length: Short, Long) 4-between (Group) repeated measures ANOVA. The division of sequences into short vs. long was based on the findings of two previous studies. According to Experiment 3 in the study by Jefferies, Matthew, and Baddeley (2004), what is called a word span for stories[4] was 15.4 (SD 2.6). Further, Isolahti's (2014) corpus-based research revealed that the mean speech sequence length of court interpreters in the consecutive mode was 10.73. Thus, the normal length of speech sequences allowing a reasonable recall of coherent prose seems to lie between 11 to 17 words.

The results revealed a significant main effect of sequence length ($F(1,86) = 117.00$, $p < 0.001$, $\eta_p^2 = 0.576$), showing that sequences shorter than 11 words were recalled better than sequences longer than 17 words by all groups. There was also a significant main effect of Group ($F(3,86) = 8395,25$, $p < 0.001$, $\eta_p^2 = 0.253$), and a significant interaction between Length × Group ($F(3,86) = 4.613$, $p = 0.005$, $\eta_p^2 = 0.139$).

Table 3. Mean scores for idea units recalled out of maximum possible in the prose recall test by group, including differences between short and long speech sequences.

Prose recall	Simultaneous interpreters $n = 21$	Consecutive interpreters $n = 21$	Foreign language teachers $n = 24$	Non-linguistic experts $n = 24$	Total $n = 90$
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Idea units for whole text (points out of maximum possible; 62)	0.84 (0.09)	0.81 (0.08)	0.77 (0.07)	0.68 (0.12)	0.77 (0.11)
ANOVA between-group					<0.001
Bonferroni corrected contrasts	<0.004	<0.004	<0.004		
Idea units for short speech sequences (≤ 11 words: Seq. 2, 8, and 11) (points out of maximum possible; 12)	0.94 (0.09)	0.93 (0.07)	0.89 (0.12)	0.87 (0.13)	0.90 (0.11)
Idea units for long speech sequences (≥ 17 words, Seq. 1,3−7, and 9−10) (points out of maximum possible; 50)	0.83 (0.11)	0.80 (0.10)	0.77 (0.08)	0.64 (0.13)	0.76 (0.13)
ANOVA between-group x speech sequence length					<0.001
Bonferroni corrected contrasts	<0.004	<0.004	<0.05		

The Bonferroni corrected contrast analysis demonstrated that the significant Group effect was due to all three language expert groups outperforming the non-linguistic experts in long speech sequences, with Bonferroni corrected significance values of $p < 0.004$ for both interpreter groups, and $p < 0.05$ for teachers. See Table 3 for more details. This indicates that all foreign language groups performed better than the non-linguistic group, especially as regards the recall of long speech sequences, but the difference between the two interpreter groups and the teachers did not reach statistical significance.

Discussion

All three foreign language expert groups outperformed the non-linguistic expert group in prose recall, and especially in the recall of long speech sequences (17 to 23 words). The result can be compared to the reading span[1] test results by Christoffels et al. (2006), in which simultaneous interpreters outperformed teachers. In our test, however, the difference between interpreters and teachers was not significant. This could be attributable to differences between the reading span test and our test: it is probable that the recall of coherent prose as opposed to recalling the last words of unrelated sentences involves memory processes which all foreign language expert groups have rehearsed equally at work. For example, the results by Jefferies et al. (2004; Exp. 3) seem to point to coherence advantage: in their research, the word span for individual words was 5.8 (SD 0.85), but for unrelated sentences and for short coherent stories 15.4 (SD 2.60).[4] Maintaining coherence in comprehension can be achieved by constructing a mental model, such as a situation model (Kintsch, 1972; Radvansky, Copeland, & Zwaan, 2005; Zwaan & Radvansky, 1998; see also Ericsson & Kintsch, 1991).

However, as teachers performed slightly less well than interpreters even in our experiment, more research is needed to find out what memory processes could be involved. Perhaps the same test design as in the present study, but with longer speech sequences (20 words or more), could reveal

Table 4. p values from Bonferroni corrected contrast analysis in between-group MANOVA for idea unit scores by sequence and by group (foreign language groups compared to non-linguistic group).

Idea unit scores by speech sequence (code: number of sequence_number of words_ maximum score)	Simulta-neous interpreter $n = 21$	Conse-cutive interpreter $n = 21$	Foreign language teachers $n = 24$	Total $n = 90$
	p	p	p	P
1st speech sequence (1_17_6)	< 0.003	< 0.003	0.017	< 0.001
2nd speech sequence (2_9_2)	1.0	1.0	0.690	0.189
3rd speech sequence (3_20_4)	0.006	< 0.003	0.006	< 0.002
4th speech sequence (4_22_4)	0.027	0.027	0.027	< 0.05
5th speech sequence (5_23_10)	0.330	0.015	0.195	< 0.05
6th speech sequence (6_20_6)	0.507	0.033	0.657	0.086
7th speech sequence (7_22_6)	0.213	0.057	1.0	0.077
8th speech sequence (8_8_4)	1.0	1.0	0.987	0.757
9th speech sequence (9_22_8)	0.027	< 0.003	0.174	< 0.005
10th speech sequence (10_23_6)	< 0.003	< 0.003	0.003	< 0.001
11th speech sequence (11_11_6)	0.126	0.006	1.0	< 0.05

significant differences not only between interpreters and teachers, but also between the two interpreter groups. The findings by Isolahti (2014), however, seem to point to sizeable individual differences between court interpreters as regards preferred speech sequence length. While the mean sequence length in her study was 10.73 words for the consecutive mode, one of the interpreters coped with lengths of 18.61 words on average, but could tolerate even lengths of 51 words or more without interrupting the speaker. Despite this, the interpreter did not lose more of the key information than the other five interpreters studied. See also General discussion for additional aspects.

Results, qualitative analysis

For the qualitative analysis, only speech sequences showing statistically significant differences between the groups were selected: Seqs. 1, 3, 4, 5, 9, 10, and 11 (see Table 4). It is noteworthy that these are not the longest sequences mentioned in the quantitative analysis (cf. Table 3). For instance, Seq. 11 with only 11 words is included, but not Seq. 9 with 22 words. Thus, it is not only the speech sequence length that seems to determine the between-group differences, but other factors are also involved (see Appendix 2). In addition, in each of the sequences listed above, only expressions (in most cases idea units) with the highest between-group differences were selected: the number of individuals recalling the expression correctly or almost correctly in the best group was to be at least twice as high as the number of individuals in the weakest group.

A closer study of the expressions which the foreign language experts recalled better than the non-linguistic experts revealed that not all of them could be classified according to traditional syntactic or other linguistic features. Therefore, a less formal classification was used. (See also the classification of memory functions in text comprehension by Ericsson & Kintsch, 1991.)

Consequently, the expressions were divided into groups as follows (for percentual sizes of between-group differences as well as the expressions with English translations, see Appendix 2):

1. Exact expressions of time (Seq. 1)
2. Expressions of emotions, essential for the message:
    - Seq. 3
    - In addition, in Seq. 5: <u>simultaneous</u> interpreters and <u>teachers</u> outperformed consecutive interpreters: 71 and 75 vs. 38%.

3. Causal relations:
    - Seq. 5
    - In addition, in Seq. 11: <u>both</u> interpreter groups succeeded better than teachers: 71 and 48 vs. 24%.

Furthermore, essential differences were identified <u>even</u> between the <u>foreign language</u> groups as follows:

4. Predicate expression:
    - Seq. 1 (<u>both</u> interpreter groups outperformed teachers: 76 and 62 vs. 38%)
5. Perfectly or almost perfectly recalled list of words:
    - Seq. 3 (<u>simultaneous</u> interpreters outperformed both other foreign language groups: 76 vs. 48 and 38%)
6. Inclusion of topic sentences:
    - Seq. 5 (<u>both</u> interpreter groups did better than teachers: 48 vs. 21%)
    - Seq. 9 (<u>both</u> interpreter groups did better than teachers: 87 and 63 vs. 42%)

In sum, the differences in prose recall between the three foreign language groups and the non-linguistic group seem to concentrate on the accuracy of recalling expressions of time and emotional content, as well as causality (Seqs. 1, 3, and 5). With one exception, though: the emotional expression in Seq. 5 was also recalled poorly by the consecutive interpreters.

Besides differences in comparison to the non-linguistic group, differences could be detected even between the foreign language groups: both interpreter groups recalled a predicate expression (Seq. 1) better than the teachers did, and they also included topic sentences in both cases in Seqs. 5 and 9. Furthermore, the simultaneous interpreters outperformed the other two foreign language groups in the accurate recall of a list containing similar words (Seq. 3) and of causality (Seq. 11).

Discussion

Qualitative analyses of the prose recall results revealed that in addition to sequence length, the type of expression seems to explain some of the differences between the foreign language and non-linguistic groups. The non-linguistic experts seemed to have particular difficulties in recalling accurately expressions of time and emotional content, as well as causality. Recalling numerical expressions, such as the expression of time in this experiment, has been shown to demand special effort (Mazza, 2001; Seleskovitch, 1976), and even experienced interpreters make errors and omissions in recalling them (Gile, 1997, 2008). As to causal relations, they form an essential element of prose comprehension (see Introduction), and because of the once-for-all character of listening, they probably play an even more important role in the comprehension of spoken prose. Furthermore, if a causal relation is expressed with a participial phrase (as in Seq. 11), listeners often have to deconstruct it in their mind into a relative clause to understand it correctly.[5] However, because of their long-time expertise in languages, native and foreign, the foreign language groups perhaps have more practice in listening to even the more demanding participial phrases, resulting in a better recall.

As to the differences detected even between the foreign language groups, the accuracy in the recall of a predicate expression (Seq. 1) of both interpreter groups compared to teachers could be related to the careful listening which is essential in the interpreters' profession. Previous research has demonstrated that topic is used for the integration of text (Lorch, 1993), and advance provision the topic or title of the prose to be recalled enhances recall (Dooling & Lachman, 1971; Einstein, McDaniel, Bowers, & Stevens, 1984; Schwarz & Flammer, 1981). As no topic was given in our experiment, recalling the first sequence containing 17 words and a rare predicate expression ('vakiinnutti asemansa' – *established its position*) apparently demanded special effort while listening.

Topic sentences exhibited another type of expressions recalled better by both interpreter groups than by the teachers. Usually, topic sentences bring additional information on a more general level, as a prelude to more essential or detailed information. So, listening to topic sentences could be more passive, as if waiting for the more essential information, thus leading to encoding processes with less attention and possibly less elaboration.[6] However, through the demand for accuracy in their work, interpreters are perhaps more practised in focusing even on topic sentences than teachers are. Topic sentences probably help interpreters to construct a mental model (situational model; see, e.g. Ericsson & Kintsch, 1991).

However, another explanation is also possible. In their experiments designed to find out whether the order of presentation affects the reading and later recall of simple two-clause arguments, Britt and Larson (2003) found that claim-first arguments were recalled more accurately than reason-first arguments. Possibly, through their work, interpreters are more apt than other experts to process topic sentences much in the same way as claims were processed by the participants in Britt and Larson's (2003) experiment: topic sentences, too, can hold important information.

In the present study, simultaneous interpreters outperformed the other foreign language groups (as well as the non-linguistic group) in the more accurate recall of a list containing similar words in Seq. 3, and in the recall of a demanding causal relation in Seq. 11. We suppose that again, the explanation is careful listening, in which simultaneous interpreters are well practised, because of the extreme time pressure while interpreting. The accurate recall of a word list also corroborates the findings of the free recall test with the same participants reported by Hiltunen et al. (2016), thus indicating that simultaneous interpreters in particular are capable of careful listening.

As to the causal relation in Seq. 11, the fact that depression, too, can affect the feeling of loneliness was mentioned in this sequence for the first time, which could have come as a surprise to the listeners resulting, in a reduced accuracy of listening. At work, however, simultaneous interpreters have to be alert all the time: if they miss some information while listening to the beginning of a sentence they risk misunderstanding the whole sentence or even a longer passage.

All these between-group explanations, however, have to be treated with particular reserve. As only one or two expressions of each type were included in the prose passage used in the present experiment, the conclusions and their possible explanations need to be confirmed by further studies.

General discussion

In prose recall, and especially in the recall of long speech sequences (17 to 23 words), all three foreign language expert groups outperformed the non-linguistic expert group. The results are consistent with the findings by Shimamura et al. (1995, see Introduction). Possibly, compared to standard population such as that in the experiment by Shimamura et al., foreign language experts, and especially interpreters, through their education and work, are practised in listening to, comprehending and remembering abstract or scientific information in much the same way as professors are.

However, the finding that the foreign language experts outperformed the non-linguistic experts in recalling long speech sequences can also, at least partly, be explained by greater familiarity with structures that are complex or otherwise demand great accuracy. Special difficulties for the non-linguistic experts were caused, among others, by the accurate recall of expressions of time and emotional content, as well as causality.

Furthermore, in the present prose recall experiment interpreters also outperformed foreign language teachers in recalling a word list containing similar words, a demanding predicate expression and a causal relation, as well as in including even the topic sentences during recall. This could be explained by an effortful concentration on listening, which has been illustrated in previous research (see, e.g. Gile, 1997, 2008). Possibly, interpreters are accustomed to listening in a different way than other people whose interest in listening is in extracting useful information.

In contrast, to be able to convey as much as possible of the information presented to their listeners, interpreters seem to pay attention even to more implicit material, such as information guiding to essential points, inferences and other structural details employed by the original speaker so as to enhance comprehension (see, e.g. Setton, 1998). The results of the present experiment seem to indicate that these same automatized methods of careful listening are transferred to situations outside the interpreting context (for additional aspects of interpreter advantage hypotheses, see García, 2014).

The findings by Dillinger (1994) appear to confirm this conclusion. If comprehension, an important subskill of prose recall, is not a specialized skill for interpreters, then the causes of the better recall of the more demanding expressions by the interpreters in this study can depend on other subskills, such as careful listening, encoding and recalling details.

In contrast, the findings by Köpke and Nespoulous (2006) do not corroborate with our results. These scholars found out that the listening span was somewhat higher for the interpreting students than for the professional interpreters. A possible explanation to these intriguing results could lie in differences between the two memory measures: listening span and prose recall. According to Daneman and Carpenter (1980) for pronoun reference questions the correlation between listening span and listening comprehension is high ($r = 0.85$) but for fact questions only moderate ($r = 0.47*$). In the present prose recall test, however, most of the expressions indicating careful listening by professional interpreters (see Appendix 2) show an efficient recall of facts, such as precise expressions of time, emotions and causal relations. It thus seems that, at least for interpreters, prose recall is a better predictor of accurate listening of facts than listening span. On the other hand, as the finding by Köpke and Nespoulous (2006) is the only one concerning the listening span of professional interpreters so far, and only a few expressions of each type were included in the present prose recall study, more research is needed to reveal what kind of memory-testing measures are best suited for testing the memory of interpreters. For the listening span test, the test design could also be modified in line with Conway et al. (2005)[7] because such modification could be especially suitable for testing interpreters' memory.

In any case, the findings of the present study are quite tentative. The length of the prose passage is limited and there are only a few examples of each of the expressions producing differences in recall between the groups. Hence, more research is needed, perhaps with prose passages designed especially to contain some of the expressions which seem to be difficult to recall for one of the foreign language groups, but less difficult for others. Another possibility could be to analyse recorded authentic productions of interpreters working in the two modes, consecutive and simultaneous, especially for these expressions.

Conclusions

In prose recall, and especially in recalling long speech sequences (17 to 23 words), all three foreign language expert groups outperformed the non-linguistic expert group. These findings seem to indicate better working memory skills of the foreign language experts, here measured with a prose recall test. Similar findings were indicated by a free recall test of unrelated words reported earlier by Hiltunen et al. (2016).

In addition, the qualitative analysis of idea units recalled better by the foreign language groups demonstrated that the recall of expressions of time, topic sentences, and demanding emotional and causal expressions requires special accuracy while listening, sometimes leading to low percentages in recall even by one or two of the foreign language expert groups, although taken as a whole, they appear to do better than the non-linguistic group. This finding seems to indicate that operating in two or more languages on a daily basis leads to some specialized cognitive skills in recall accuracy for the foreign language experts. Such cognitive skills can, however, be expertise-dependent. For exam-

ple, interpreters seem to listen to spoken prose with an effortful concentration, noticing even pragmatic or structural details which a person with some other expertise can overlook completely.

Notes

1. In the reading span test, the participant has to read visually presented unrelated sentences and subsequently to recall the last word of each sentence in serial order.
2. Immediate vs. delayed recall – In immediate recall the items presented are recalled immediately after presentation (see also note 4). In delayed recall the items are recalled after a delay of several minutes or hours up to 24 h or more.
3. The average length of Finnish words is high, leading to speech rates with fewer words per minute than in English, for example, where 150 to 160 words per minute is recommended for audiobooks, for instance (Williams, 1998).
4. Memory span refers to the number of items that can be repeated immediately in correct order for 50% of the time (Neath and Surprenant, 2003, p. 464). If the items are words, the memory span is called a word span; if they are sentences, it is called a sentence span. The sentences can be unrelated or form a coherent story. To allow a better comparison between the two, in the study by Jefferies et al. (2004) the expression 'word span' was used both for individual words and for sentences; as a result, the maximum number of words correctly recalled was counted for both the word span and the sentence span measures.
5. It is noteworthy that according to the recommendations for writing plain Finnish, participial phrases should be avoided by replacing them with relative clauses much simpler to understand (see, e.g. Kulkki-Nieminen, 2010, for a review).
6. Elaboration refers to a conscious, explicit effort of finding associations to the stimuli to be remembered, in contrast to other encoding processes, such as passive listening or reading, which lead to more implicit learning. Elaboration has been shown to enhance memory in many different tests; see, e.g. Craik and Lockhart (1972), Hyde and Jenkins (1973).
7. The modification of the reading span (and listening span) test by Conway et al. (2005) is twofold. Firstly, instead of the last words to be recalled, an arbitrary word is added at the end of each sentence to test recall. Secondly, instead of blocks of sentences increasing gradually from two sentences to seven, the size of blocks is random. The first modification would be better suited to testing foreign language experts, for two reasons. First, depending on the language, the final words of a sentence can represent different parts of speech and thus be more difficult to recall in some languages than in others. Second, sentence coherence, which can also differ between languages, would have no impact on the recall of arbitrary words. Also, the second modification would remove the possibility of easing the memory load through anticipation of the number of blocks, and as interpreters are especially practised in anticipation while interpreting, this could give a special advantage to them as compared to other expert groups.

References

Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Clarendon Press.
Baddeley, A. D. (1990). *Human memory. Theory and practise*. Revised edition (1997). Hove, UK: Psychology Press.
Baddeley, A. D., Logie, R., & Nimmo-Smith, I. (1985). Components of fluent reading. *Journal of Memory and Language*, *24*, 119–134.
Blunt, J. R., & Karpicke, J. D. (2014) Learning with retrieval-based concept mapping. *Journal of Educational Psychology*, *106*(3), 849–858.

Britt, M. A., & Larson, A. A. (2003). Constructing representations of arguments. *Journal of Memory and Language*, *48*(4), 794–810.

Cavallini, E., Cornoldi, C., & Vecchi, T. (2009). The effects of age and professional expertise on working memory performance. *Applied Cognitive Psychology*, *23*, 382–295.

Chi, M. T. H. (2003). Two approaches to the study of expert's characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21–30). Cambridge, UK: Cambridge University Press.

Christoffels, I. K., de Groot, A. M. B., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language*, *54*, 324–345.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 763–786.

Corgiat, M. D., Templer, D. I., & Newell, T. C. (1989). The effects of presentation modality, task demand, and content structure on age-related memory differences for prose. *International Journal of Aging & Human Development*, *29*(1), 53–65.

Cornoldi, C., & De Beni, R. (1991). Memory for discourse: Loci mnemonics and the oral presentation effect. *Applied Cognitive Psychology*, *5*, 511–518.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Behavior*, *11*, 671–684.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*, 422–433.

Dillinger, M. (1994). Comprehension during interpreting: What do interpreters know that bilinguals don't? *The Interpreters' Newsletter*, *3*, 41–58.

Dooling, D. J., & Lachman, R. (1971). Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, *88*, 216–222.

Einstein, G. O., McDaniel, M. A., Bowers, C. A., & Stevens, D. T. (1984). Memory for prose: The influence of relational and proposition-specific processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 133–143.

Ericsson, K.A., & Charness, N. (1994). Expert performance. Its structure and acquisition. *American Psychologist*, *49*(8), 725–747.

Ericsson, K. A, & Kintsch, W. (1991). Memory in comprehension and problem solving: A Long-term working memory. *Publication Number 91–13*, Institute of Cognitive Science, University of Colorado, Boulder: CO.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102:2*, 211–245.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273–305.

Ericsson, K. A., Krampe, R. Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 361–406.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits in the empirical study of expertise. In K. A. Ericsson, & J. Smith (Eds.), *Toward a general theory of expertise: prospects and limits* (pp. 1–38). Cambridge: Cambridge University Press.

Ericsson, K. A., & Williams, A. M. (2007). Capturing naturally occurring superior performance in the laboratory: Translational research on expert performance. *Journal of Experimental Psychology: Applied*, *13*(3), 115–123.

Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and Language*, *27*, 235–244.

Fletcher, C. R., Hummel, J. E., & Marsolek, C. J. (1990). Causality and the allocation of attention during comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(2), 233–240.

García, A. M. (2014). The interpreter advantage hypothesis. Preliminary data patterns and empirically motivated questions. *Translation and Interpreting Studies 9*(2), 219–238.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 430–445.

Gile, D. (1997). Conference interpreting as a cognitive management problem. In J. H. Danks, G. M. Shreve, S. B. Fountain, & M. K. McBeath (Eds.), *Cognitive processes in translation and interpreting* (pp. 196–214). Thousand Oaks, CA: Sage Publications.

Gile, D. (2008). Local cognitive load in simultaneous interpreting and its implications for empirical research. *Forum*, *6*(2), 59–77.

Glover, J. A. (1981). Distinctiveness of encoding: The effects of paraphrasing and drawing inferences on memory from prose. *Journal of Educational Psychology*, *73*(5), 736–744.

Graesser, A .C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395.

Hiltunen, S., Vik, G.-V., Pääkkönen, R., & Krause, Chr.M. (2016). On interpreters' working memory and executive control. *International Journal of Bilingualism, 20*(3), 297–314.

Hoffman, R. R. (1998). How can expertise be defined? Implications of research from cognitive psychology. In R. Williams, W. Faulkner, & J. Fleck (Eds.), *Exploring expertise* (pp. 81–100). New York, NY: Macmillan.

Hyde, T. S., & Jenkins, J. J. (1973). Recall of words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Behavior*, *12*, 471–480.

Isolahti, N. (2014). Tulkkauksen tarkkuus rikosoikeudenkäynnissä – saavuttamaton ihanne. Akateeminen väitöskirja, Tampereen yliopisto, Acta Universitatis Tamperensis 1892 (print); Acta Electronica Universitatis Tamperensis 1374 (pdf.) Tampere: Suomen Yliopistopaino Oy – Juvenes Print.

Jackson, M. D., & McClelland, J. L. (1979). Processing determinants of reading speed. *Journal of Experimental Psychology: General*, *108*(2), 151–181.

Jefferies, E., Matthew, A. L. R., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language*, *51*, 623–643.

Jiang, L., Han, Ch., Jiang, J., & Feng, Y. (2014). The sociological turn in the interpreter's role. Discourse interpreting filters. *Translation and Interpreting Studies*, *9*(2), 274–298.

Kalakoski, V., & Saariluoma, P. (2001). Taxi drivers' exceptional memory of street names. *Memory & Cognition*, *29*(4), 634–638.

Kintsch, W. (1972). How readers construct situation models for stories: The role of syntactic cues and causal inferences. In S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes. Essays in honor of William K. Estes* (pp. 261–278). Hillsdale, NJ: Erlbaum.

Köpke, B., & Nespoulous, J. (2006). Working memory in expert and novice interpreters, *Interpreting*, *8*(1), 1–23.

Kulkki-Nieminen, A. (2010). Selkoistettu uutinen. Lingvistinen analyysi selkotekstin erityispiirteistä. Akateeminen väitöskirja. Dissertation, Tampere University.

Lorch, R. F., Jr.., (1993). Integration of topic and subordinate information during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1071–1081.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, *132*, 199–207.

Macaruso, P., & Shankweiler, D. (2010). Expanding the simple view of reading in accounting for reading skills in community college students. *Reading Psychology*, *31*, 454–471.

Macnamara, B. N., Moore, A. B., Kegl, J. A., & Conway, A. R. A. (2011). Domain-general cognitive abilities and simultaneous interpreting skill. *Interpreting*, *13*(1), 121–142.

Marschark, M., De Beni, R., Polazzo, M. G., & Cornoldi, C. (1993). Deaf and hard of hearing adolescents' memory for concrete and abstract prose: Effects of relational and distinctive information. *American Annals of the Deaf*, *138*(1), 31–39.

Mazza, Chr. (2001). Numbers in simultaneous interpreting. *The Interpreters' Newsletter*, *11*, Trieste, EUT Edizioni Università di Trieste, 87–104.

McDaniel, M. A. (1984). The role of elaborative and schema processes in story memory. *Memory & Cognition*, *12*(1), 46–51.

McDaniel, M. A., & Kerwin, M. L. E. (1987). Long-term prose retention: Is an organizational schema sufficient? *Discourse Processes*, *10*, 237–252.

Miangah, T. M. (2014). Effect of prior Knowledge on recall and comprehension. Retrieved from http://mosavit.ir/images/articlesB/6_prior_knowledge_on_recall_and_comprehension.pdf (accessed 15.01.2015).

Mills, C. B., Diehl, V. A., Birkmire, D. P., & Mou, L. (1993). Procedural text: Predictions of importance ratings and recall by models of reading comprehension. *Discourse Processes*, 16, 279–315.

Neath, I., & Surprenant, A. M. (2003). *Human memory. An introduction to research, data, and theory* (2nd ed.) Belmont, CA: Wadsworth & Thomson Learning.

Obler, L. K. (2012). Conference interpreting as extreme language use. *International Journal of Bilingualism*, *16*(2), 177–182.

O'Brien, E. J., & Albrecht, J. E. (1991). The role of context in accessing antecedents in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 94–102.

Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Memory and Language*, *24*, 59–88.

Pöchhacker, F. (2004). *Introducing interpreting studies*. London: Routledge.

Radvansky, G. A., Copeland, D. E., & Zwaan, R. A. (2005). A novel study: Investigating the structure of narrative and autobiographical memories. *Memory*, *13*(8), 796–814.

Raitasalo, R. (2006). Pakotettu yksinäisyys, Tutkittu juttu, *Kelan sanomat 1/2006, verkkolehti*.

Schwarz, M. N. K., & Flammer, A. (1981). Text structure and title – effects on comprehension and recall. *Journal of Verbal Learning & Verbal Behavior*, *20*, 61–66.

Seleskovitch, D. (1976). Interpretation: A psychological approach to translating. In R.W. Brislin (Ed.), *Translation: Applications and research* (pp. 92–166). New York: Gardner Press.

Setton, R. (1998). Meaning assembly in simultaneous interpretation. *Interpreting*, *3*(2), 163–199.

Shimamura, A. P., Berry, J. M., Mangels, J. A., Rusting, Ch.L., & Jurica, P. J. (1995). Memory and cognitive abilities in university professors: Evidence for successful aging. *Psychological Science*, *6*(5), 271–277.

Signorelli, T. M., Haarmann, H. J., & Obler, L. K. (2012). Working memory in simultaneous interpreters: Effects of task and age, *International Journal of Bilingualism*, *16*(2), 198–212.

Sisco, S. M., Marsiske, M., Gross, A. L., & Rebok, W. (2013). The influence of cognitive training on older adults' recall for short stories, *Journal of Aging and Health*, *25*(8 suppl.), 230S–248S.

Tiselius, E. (2013). Experience and expertise in conference interpreting. An investigation of Swedish conference interpreters. Dissertation, University of Bergen, Norway.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, *24*, 612–630.

Williams, J. R. (1998). Guidelines for the use of multimedia in instruction. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, 1447–1451.

Yudes, C., Macizo, P., & Bajo, T. (2011). The influence of expertise in simultaneous interpreting on nonverbal executive processes. *Frontiers in Psychology*, *309*(2), 1–9. Retrieved from http://www.frontiersin.org/Psychology (accessed 05/07/2013).

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.

Author biographies

Sinikka Hiltunen has MA degrees both in translation and interpreting (Tampere University) and in cognitive science (Helsinki University). Over two decades, she has worked as Translator and Consecutive Interpreter, mainly for manufacturing industries. She is also an authorized translator. She is currently a doctoral student in Helsinki University and also works for various educational organizations as free-lance Lecturer, Educator and Trainer on topics related to linguistics, psychology, and cognitive sciences.

Gun-Viol Vik, PhD, senior university Lecturer at University of Vaasa, Department of Scandinavian Languages. Research interests are expertise in interpreting, professional language and communication, as well as bi- and multilingual practice in organizations. She works also as Professional Conference Interpreter.

Appendix 1

Text for prose recall

Extract from the on-line article 'Pakotettu yksinäisyys' ('*Forced Loneliness*', Raimo Raitasalo, 2006) divided into speech sequences with indication of word quantity and maximum point score for each idea unit (the critical verb phrase shaded, see Mills et al., 1993). English translation follows.

Speech sequence with code (number of sequence_number of words_maximum points). The critical word for determination of idea unit is shaded.	Number of words	Points, max.
1st speech sequence, code 1_17_6		
Yksinäisyyttä kartoittava tutkimus	17	2
joka vakiinnutti asemansa hyvinvoinnin ja sosiaalisten ongelmien tutkimusperinteessä 1970-luvun lopulla,		2
lähtee tavallisesti kolmesta perusolettamuksesta:		2
2nd speech sequence, code 2_9_2		
Se on jokaisen ihmisen omakohtainen kokemus, ei objektiivinen tila.	9	2
3rd speech sequence, code 3_20_4		
Kokemuksena se ei useinkaan ole miellyttävä, vaan kiusaava ja välillä lamaannuttava jopa henkisesti halvaannuttava.	20	2
Se on yhteydessä läheisten ihmissuhteiden tyydyttävyyteen.		2
4th speech sequence, code 4_22_4		
Pakotetusta yksinäisyyden tunteesta puhutaan silloin,	22	2
kun ihminen ei pysty terveydentilansa, asuinpaikkansa tai taloudellisen tilanteensa johdosta toteuttamaan omia sosiaalisen liittymisen ja yhteisyyden tarpeitaan.		2
5th speech sequence, code 5_23_10		
Erään kattavan hollantilaisten ikääntymisen vaikutuksia seuranneen tutkimuksen perusteella	23	2
vastaus on selvä;		2
iän kohotessa		2
vanhuudessa myös haittaavat yksinäisyyden tunteet lisääntyivät,		2
voimakkaimmin hyvin iäkkäiden keskuudessa.		2
6th speech sequence, code 6_20_6		
Yllättävältä saattaa tuntua se tutkimuksen tulos,	20	2
jonka mukaan puolisonsa kanssa yhdessä elävät kokivat		2
iän kohotessa		2
enemmän yksinäisyyttä kuin yksin asuvat.		
7th speech sequence, code 7_22_6		
Tämän katsottiin johtuvan siitä,	22	2
että iän myötä huoli puolisosta ja hänen terveydentilastaan kasvoi		2
ja esille tulleet terveysongelmat alkoivat lisätä pelkoa toisen menettämisestä.		2

Appendix 1. (Continued)

Speech sequence with code (number of sequence_number of words_maximum points). The critical word for determination of idea unit is shaded.	Number of words	Points, max.
8th speech sequence, code 8_8_4		
Samalla myös epävarmuus lisääntyi	8	2
ja yksinäisyyden tunteet voimistuivat.		2
9th speech sequence, code 9_22_8		
Ikääntyvien väliset suhteet muuttuivat,	22	2
kun hauraus lisääntyi.		2
Iän kohotessa		2
aikaisempi perheen ylläpitäjä ja kannustaja saattoikin muuttua vaativaa huolenpitoa ja hoivaa tarvitsevaksi potilaaksi.		2
10th speech sequence, code 10_23_6		
Tulos osoitti sen,	23	2
että toisen kanssa yhdessä eläminen ei enää myöhemmällä iällä, iän kohotessa suojannutkaan yksinäisyyden tunteilta		2
niin kuin se oli suojannut nuorempana.		2
11th speech sequence, code 11_11_6		
Voimakkaimmillaan masentuneisuuteen liittyvää yksinäisyyttä	11	2
koettiin silloin,		2
kun oma puoliso oli menetetty.		2

Text for prose recall, translation (by Heli Mäntyranta)

1st speech sequence

Research mapping loneliness
which established its position in the research tradition of wellbeing and social problems in the late 1970s
usually starts out from three basic assumptions:

2nd speech sequence

It is the personal experience of each individual, not an objective state.

3rd speech sequence

As an experience, it is not often pleasant, but troubling and sometimes numbing and even mentally paralysing.
It is connected to the level of satisfaction from one's intimate relationships.

4th speech sequence

A forced loneliness is spoken of
when a person is not able to, due to his or her state of health, place of residence or financial situation, fulfil his or her personal needs for social attachment and community.

5th speech sequence

On the basis of a comprehensive Dutch study that examined the effects of ageing,
the answer is clear:

with advancing age,
in older years intrusive feelings of loneliness also increased
most strongly among the very elderly.

6th speech sequence

The finding may appear surprising
that those living together with a spouse experienced
with advancing age
more loneliness than those living alone.

7th speech sequence

This was considered to be due to the fact that,
with age, concern over the spouse and his or her health increased
and the emergence of health problems began to increase the fear of losing the other.

8th speech sequence

At the same time, uncertainty also increased
and the feelings of loneliness grew stronger.

9th speech sequence

The relationships between ageing people changed
as frailty increased.
With increasing age,
the one that had previously been the upholding and encouraging one in the family could become a
patient in need of demanding care and nurturing.

10th speech sequence

The finding showed
that living together with another person did not, at a later age as age increased, protect one from
feelings of loneliness
the way it had protected when younger.

11th speech sequence

The experience of loneliness related to depression
was the strongest
when the person had lost his or her spouse.

Appendix 2

Main qualitative between-group differences in recall of the most demanding expressions in speech sequences 1, 3, 5, 9 and 11.

Speech sequence, problem type	Phrase	(Almost) Correctly recalled, per cent from the participants*				
		Simultaneous interpreters	Consecutive interpreters	Foreign language teachers	Non-linguistic experts	Mean % (SD)
1. Precise expression of time - Seq. 1	1970-luvun lopulla (*at the end* of the 1970s)	43	43	50	21	39.25 (12.6)
2. Expression of emotions essential for understanding the message: Seq. 3	yhteydessä läheisten ihmis-suhteiden tyydyttävyyteen (*is connected to the level of satisfaction from one's intimate relationships*)	38	29	29	4	25 (14.6)
- Seq. 5	haittaavat yksinäisyyden tunteet (*troubling feelings of loneliness*)	71	38	75	42	56.5 (19.2)
3. Causal relation - Seq. 5. loneliness increasing with growing age	- voimakkaimmin hyvin iäkkäiden keskuudessa (*most strongly among the very elderly*)	48	57	46	25	44 (13.5)
- Seq. 11 (perfectly correct)	Voimakkaimmillaan masentuneisuuteen liittyvää yksinäisyyttä…(*The experience of loneliness related to depression was the strongest…*)	71	48	24	14	39.25 (25.5)
4. Predicate expression - Seq. 1, (*became famous, began, is being studied since* accepted as well)	vakiinnutti asemansa (*established its position*)	76	62	38	33	52.25 (20.3)
5. Word list - Seq. 3, attributes for loneliness (for full points three out of four required)	ei miellyttävä, kiusaava, lamaannuttava, henkisesti halvaannuttava (*not pleasant, troubling, numbing, mentally paralyzing*)	76	48	38	21	47.75 (12.6)
6. Topic sentence included: - Seq. 5	…vastaus on selvä (*the answer is clear*)	48	48	21	13	32.5 (18.2)
- Seq. 9	… kun hauraus lisääntyi (*as frailty increased*)	87	62	42	37	57 (22.7)
						43.5 (18.8)