5 Using Machine-Translated Texts to Generate L3 Learners' Metalinguistic Talk

Eva Thue Vold

Introduction

Metalinguistic awareness, or “the ability to focus attention on language as an object in and of itself, to reflect upon language, and to evaluate it” (Thomas 1988, 531), is a central component of metacognition in language learning (Jessner 2018, this volume; Haukås 2018, this volume). Furthermore, it has been described as a key factor in third language (L3) learning (Jessner 2008, 270, 275; Jessner 2018, this volume). Nevertheless, L3 teachers often complain that students' metalinguistic awareness is low and that they have insufficient knowledge about grammatical terms from language arts and L2 classes (see for example, Haukås 2016, 11). Hence, there seems to be a need for L3 activities that can enhance students' metalinguistic awareness and analytical skills.

The ability to critically examine and evaluate language and language use is especially important when working with digital tools such as free online machine translation (MT). Teachers at all levels often advise against using such tools, claiming that students do not learn anything from them and that the resulting student texts are poor in quality (see Clifford, Merschel, and Munné 2013). However, this may be the result of uncritical use of the tools (see Williams 2006). Research suggests that MT can be a resource for learning, provided that students are given the opportunity to develop sufficient knowledge about language and about the tools themselves (see Jolley and Maimone 2015; Niño 2009).

This study sets out to examine what characterises students' metalinguistic talk when they discuss the quality of machine-translated texts. In an attempt to answer this overarching research question, students from a French as a third language class in an upper secondary school in Norway were given a task during which they read two machine-translated versions of the same text. The students' task was to decide which translated version was best and then explain, in as much detail as possible, why it was better than the other. The task was aimed at developing both students' metalinguistic awareness and their competence in using MT tools critically. Both are highly relevant competences for a continuing language learning process. It is the

first aspect—the students’ metalinguistic awareness—that will be addressed in this chapter.

Literature Review

Students’ Metalinguistic Awareness

There is no clear-cut distinction between metalinguistic awareness or metalinguistic knowledge,¹ on the one hand, and linguistic knowledge, on the other. However, researchers generally agree that what characterises the metalinguistic dimension is that it is explicit, declarative and verbalisable (R. Ellis 2004; Roehr 2008; both cited in Gutiérrez 2013, 148; Roehr and Gánem-Gutiérrez 2013, 4–6). Metalinguistic awareness includes aspects of linguistic knowledge, such as explicit knowledge of grammar and pronunciation rules and of word meanings, as well as the ability to analyse and explain grammatical phenomena, compare linguistic structures across languages, make use of prior linguistic knowledge, use metalinguistic terms when talking about and describing language and reflect upon language use (see also discussions in Hasselgård 2018, this volume; Jessner 2018, this volume). The term metalinguistic awareness thus covers both aspects of explicit knowledge identified by R. Ellis (2004): analysed knowledge, i.e. awareness of grammatical rules and linguistic features; and metalanguage, i.e. the ability to verbalise analysed knowledge using technical terminology.

Previous research on metalinguistic awareness and its role in language learning has shown that these two phenomena are strongly interrelated (Cenoz, Hufeisen, and Jessner 2001; Jessner 2008). However, it is not clear from the research whether metalinguistic awareness is a prerequisite for success in L3 learning in school settings or whether such learning in itself enhances students’ metalinguistic awareness (Jessner 2008, 275). In addition, as Gutiérrez (2016) points out, findings may differ depending on what components of metalinguistic knowledge are measured, how they are measured and how the constructs are defined. One of the most frequently applied procedures in this field is the examination of learners’ collaborative interaction while performing L2/L3 tasks (see Gutiérrez 2013, 149). Pair or group work is assumed to trigger interaction related to language aspects, and the assumption is that such interaction, which from a socio-cultural perspective is often referred to as collaborative dialogue (Swain 2000; Swain, Kinnear, and Steinman 2015), can facilitate language learning insofar as it engages the participants in problem solving and potentially in linguistic knowledge construction (Lyle 2008; Swain 2000). Such collaborative talk involves metatalk, i.e. talk about language use (Swain 1998). Storch (2008) studied the metatalk of university students enrolled in a university ESL course as they engaged in a collaborative text reconstruction task. The concept of metatalk was operationalised in terms of language-related episodes (LRE), i.e. episodes in which the students talked about their

own or their peers' linguistic output (see also Swain and Lapkin 2001). An analysis of learners' performance at different stages revealed that the level of student engagement in the language-related discussion that arose had positive effects on learning and consolidation outcomes. The level of engagement was highest when grammar issues were discussed. Storch (2008, 111) concluded that opportunities to verbalise and deliberate about language are beneficial for learning, but that elaborate engagement is more effective than limited engagement. It seems reasonable to assume that, at least to some extent, high levels of engagement reflect high levels of metalinguistic awareness. Thus, there seems to be a mutual, positive reinforcement between metalinguistic awareness and language learning.

In line with the prevailing idea that metalinguistic awareness plays an important role in language learning, the current curriculum for foreign languages in Norway gives the topic of metalinguistic awareness a fairly prominent role. The curriculum is structured into three main subject areas, *Language Learning*, *Communication* and *Language, Culture and Society*, with competence aims formulated for each area. Competence aims related to metalinguistic awareness are subsumed under the subject area *Language Learning* and refer, e.g. to students' ability to make use of previous knowledge of languages when learning a new one as well as their ability to examine similarities and differences between languages as part of the language learning process (Norwegian Directorate of Education and Training 2006). Despite this focus on the role of metalinguistic awareness, few studies on Norwegian L3 learners have addressed such issues. One exception is the study by Hansejordet (2009), who studied lower secondary students' use of metalinguistic terminology in a Spanish L3 class. The findings revealed that students use very few grammatical terms when talking about language—they tend to stick to everyday terms. Sometimes this works well in the situation, whereas other times the students' lack of understanding of underlying grammatical structures leads to confusion and impedes communication within the group (Hansejordet 2009, 32–3, 83). Hansejordet's findings mirror Swain's (1998) view that metatalk does not necessarily involve the use of technical terms. A case in point is Gutiérrez' (2013) study of Spanish university students' metalinguistic knowledge, using an immediate recall technique in a collaborative text production task. Every time a covert metalinguistic activity was identified, meaning an episode in which the students commented on the language use (for instance, proposed a change in the text) without giving any explicit expression of metalinguistic thinking, the researcher prompted the student to give reasons for his or her suggestion. It was found that in almost 80% of the cases, metalinguistic knowledge turned out to be the motivation for the student's comment. Gutiérrez' (2013, 163) hypothesis is that students may not always verbalise their underlying metalinguistic knowledge representations because they either do not have the adequate resources or do not see the need to do so, as they assume that their fellow students share the same knowledge. In a later study, Gutiérrez (2016)

found that analysed knowledge (see above) was more strongly correlated with language learning than was metalanguage. He concludes that knowledge of technical terminology is not as important in the language learning process as the understanding of grammatical features. Nevertheless, studies have shown that more proficient users tend to use more metalanguage, and advanced learners tend to focus more than intermediate learners on metalinguistic aspects (Fortune 2013).

Research from European countries shows that multilingual teachers (i.e. teachers who master at least three languages) display higher metalinguistic awareness than bilingual teachers (E. Ellis 2004; Otwinowska 2014), but they may lack strategies for applying this awareness in the language-learning classroom (Otwinowska 2014). Research from the Norwegian context reveals similar tendencies: Many L3 teachers feel unsure about how to approach the main subject area of language learning in the curriculum, under which metalinguistic awareness and other types of metacognitive competence, such as self-assessment and critical use of digital tools, are subsumed. In a study of 145 language teachers in Norwegian upper secondary schools, L2 as well as L3 teachers reported that they rarely worked on enhancing learners' language learning awareness (Haukås 2012; see also Haukås 2018, this volume). Similarly, based on a survey of 85 Norwegian lower secondary L3 teachers combined with in-depth small-scale observational studies and interviews, Heimark (2013, 214) concludes that teachers seem to have few reflections related to a metaperspective on language learning compared with their reflections on other aspects of the curriculum, such as communicative skills and cultural content knowledge. Likewise, a quantitative survey of both newly qualified and experienced L3 teachers, in which the teachers were asked about their self-perceived competence for teaching different components of the subject, reveals that the teachers felt uncertain about the teaching of language learning awareness in general and of metalinguistic awareness in particular. The teachers felt most insecure about the subcomponent of metalinguistic awareness, on par only with the component of teaching self-assessment (Vold 2017). Many studies point to the absence of plurilingual training in teacher education to explain why even multilingual teachers often seem to ignore learners' language backgrounds and rarely use activities involving cross-linguistic comparisons in their teaching (De Angelis 2011; Otwinowska 2014). Cross-linguistic comparisons based on machine translation are one type of activity that teachers could be introduced to during their teacher training and could subsequently apply in their classrooms to enhance learners' awareness of the similarities and differences between languages.

Machine Translation as a Language Learning Tool

Most research on MT stems from the fields of computational linguistics and translation studies. Whereas the former focuses on technical aspects, the

latter often focuses on the role of MT in courses for trainee translators and addresses questions such as post-editing, pre-editing and quality assessment (see, for example, Gaspari, Almaghout, and Doherty 2015; Niño 2008). Although these types of questions can inform our understanding of the use of MT tools by language learners, research that specifically addresses the role of MT in language learning is scarce. For example, in Golonka et al.'s (2014) thorough review of the effectiveness of technologies in foreign language learning, MT is not included. This is understandable, as MT was not originally created for language learning purposes (Somers 2001, 28) and hence was not developed as a computer assisted language learning (CALL) tool. On the other hand, the lack of literature on the topic is somewhat surprising since a large amount of anecdotal evidence from language teachers and students indicates that language learners in school contexts make extensive use of free online machine translation (see Garcia and Pena 2011, 472; Williams 2006, 566). As these tools steadily improve, this trend is bound to increase (Garcia and Pena 2011, 472). Thus, according to Garcia and Pena (2011, 472) and Clifford, Merschel, and Munné (2013), the use of free online machine translation for language learning is a topic that can no longer be ignored by the research field.

The few research papers that address this issue reveal that the use of MT programmes for language learning is controversial. Steding (2009) adopts a critical stance, fearing that with the constant development of better software, students would “rely on [MT] more and more and not develop their own skill in the language” (Steding 2009, 178), consequently concluding that “our first goal must be to prevent the use of MT” (Steding 2009, 188). Somers (2001) also expresses reluctance, stating that since MT software is “generally not designed with language learners in mind, [. . .] one should be a little wary of using it for this purpose” (Somers 2001, 28). In addition, Somers, Gaspari, and Niño (2006, 42) assert that the MT techniques used for learning purposes “carry with them the danger that students, particularly beginners, cannot readily identify examples of bad usage, and have a not necessarily justified ‘trust’ in the accuracy of computer output”. On the other hand, the few (quasi-)experimental studies that exist show some positive effects. Niño (2008) performed a comparative analysis of error frequency in the translations produced by two groups of advanced students of Spanish: One group did post-editing of machine-translated texts, whereas the other group translated the same texts on their own. She concludes that, for advanced students, MT post-editing was beneficial in several ways, and it was “especially good for [. . .] raising language awareness through error correction and detection” (Niño 2008, 44–5). Niño’s students were adult advanced university students enrolled in a course on MT, and they were thus not comparable to a group of more or less motivated secondary school students learning an L3. In contrast, Garcia and Pena’s (2011) study involved beginners. Using an experimental design, they compared student writings with and without the help of MT and found that MT helps beginners to

communicate more (as measured by the number of words produced) and better (as measured by blind marking). However, judging from screen recordings showing the number of pauses and the number of successful and unsuccessful editing interventions, they also conclude that beginners probably learn more from writing directly in the target language using their own words (Garcia and Pena 2011). If so, this finding echoes Steding's (2009) concern that although students can rely on this kind of tool for communication or school assignment purposes, it will not help them to develop language skills.

Since MT is undoubtedly here to stay, the research field as a whole and language teachers specifically need to focus on how it can be used for learning purposes and not solely for communication. In this vein, Williams (2006) advocates the use of MT for developing electronic literacy and language awareness. In his view, students' misuse of MT stems from the fact that they use such tools uncritically. They need to develop a critical literacy towards technology as well as the ability to evaluate digital tools (Williams 2006, 567). Although Williams's paper includes interesting reflections and good advice on how to use MT for learning purposes, it does not report from classrooms in which MT was actually used. Niño (2009) also gives an overview and evaluation of different ways of using MT for learning purposes, but her suggestions seem to have been applied only with university-level language learners or trainee translators. In sum, the literature on foreign language teaching and learning provides little evidence that the use of MT can be beneficial for language learners in a school context.

The Current Study

Aim

In order to address, with empirical evidence, the question of how and whether the use of MT can be beneficial for language learning, the present small-scale study implemented some of the suggestions from the literature for using MT for learning purposes in a school context. It set out to examine what characterises students' metalinguistic talk when they discuss the quality of machine-translated texts. The students worked in small groups based on their relative achievement levels in French. The discussions between the students were analysed with regard to the following research questions (RQ):

- RQ1: How do the students talk about the texts, and to what extent do they use grammatical terms?
- RQ2: What kinds of errors do they talk about?
- RQ3: What differences are there between the different achievement levels?

Participants

The research project was carried out in a French as a foreign language class in an upper secondary school in the urban surroundings of Oslo. The students, all around 17 years of age, were in their second year of a general studies course. They had studied French as an L3 for three years in lower secondary school and, by the time the project was undertaken, for 1.5 years in upper secondary school. At this stage, students have only half a year left of the obligatory L3 course, and thus they are among the most advanced L3 students in the Norwegian secondary school system. This level was chosen because the research literature and teacher experiences indicate that, to be able to evaluate target language MT texts, students need a certain level of linguistic competence in the target language as well as some translation skills (see Niño 2009). Although the participating students were in their last year of the obligatory L3 course, their L3 level was not advanced. Few teaching hours and a lack of L3 input make it difficult for Norwegian L3 students to reach beyond the B1 level of the CEFR. The true proficiency level of the participating students was not known, but, according to the class teacher, many students in this group were weaker than they should have been at this level. Their level of motivation varied; some were not motivated to learn an L3, while others liked the subject and wanted to learn.

The class consisted of 30 students, all of whom gave their written consent to participate in the project. On the day that the project was to be carried out, seven students were absent from the L3 class. The participant group thus consisted of 23 students: 13 girls and 10 boys.

The participating teacher was an experienced French instructor. Collaborative preparations were mostly done by email, but a preparatory meeting was also held immediately before the lesson in which the project was conducted.

Materials: The Source Text

A source text in Norwegian was chosen in collaboration with the teacher. The class was working with geography and was planning an excursion to Nice, so a text about Nice from the website *Reisdit.no* was chosen. The text was informative, factual and tourist-oriented, 320 words long and not particularly difficult compared with regular French language textbooks for this level. A few typing errors in the source text were corrected before the text was translated into French by Google Translate and Bing translator.

A pre-analysis of the machine-translated texts revealed that the translation programmes made several errors of different types that students at this level could be expected to detect. Among the most salient were morphological errors, lexical errors and syntactic errors. Below are some examples of the errors the programmes made in each category.

Morphological errors comprise different types of agreement errors (1–3), verb conjugation errors (3) and omissions of function words (3–4).²

1. *Nice est une des villes les plus visités (GT)
2. *Une petite pharmacies (BT)

In the first example, the noun (*villes*) is in the feminine plural, while the adjective (*visités*) is in the masculine plural. In the second example, the noun (*pharmacies*) is in the feminine plural whereas the adjective (*petite*) and the indefinite article (*une*) are in the feminine singular.

3. *17ème Cathédrale de Ste-Réparate valent un visite (BT)

In this example, the verb (*valent*) is in the plural, while the noun (*cathédrale*) is in the singular. There is also an error in determiner—noun agreement: *un* is the masculine indefinite article, while *visite* is a feminine noun. In addition, there is an omission of the definite article in the first noun phrase.

4. *Vous avez tout simplement pas été à Nice (BT)

Here it is the negation particle *ne* that has been omitted—in written French, this particle is obligatory and is placed before the finite verb.

Lexical errors include untranslated words as well as erroneous choices of lexical items, which in turn lead to wrong or nonsense semantic meanings.

The Norwegian words *hit* (English *here*) and *dump* (English *dump*) are left untranslated by Google, while *illusjonistiske* (English *illusionistic*) is not translated by Bing.

As for the choice of lexical items, this seems to be especially challenging when the programmes have to deal with polysemous words. For example, Bing translates *travel og livlig by* (busy and lively town) as *une ville occupée et animée*, where *occupée* would be understood as invaded or conquered rather than busy. Another example is when Bing translates *lett rødvin* (light red wine) as *vin rouge facile*. *Facile*, meaning “easy/uncomplicated”, cannot be used in this context. Google, on the other hand, correctly chose the word *léger* (light / not heavy).

Another striking example of the erroneous choice of semantic meaning is when the name of the town, Nice, is not recognised as a proper noun and is thus translated by Google as *agréable* (English *nice/pleasant*) and by Bing as *belle* (English *nice/beautiful*). Both programmes obviously use English when they do not identify the word as Norwegian. Phrasal verbs are also hard to render; the Norwegian expression *dump ned* (sit down) becomes the barely comprehensible *décharge vers le bas* (literally *discharge downwards*) in the Bing version.

Syntactic errors refer to errors in word order and sentence structure. Both programmes have problems with this, and the longer and more complicated

the clause or phrase, the worse the result. Compound nouns in the source text often trigger inaccurate word orders in the translation. For example, Bing renders *eksklusive lunsjrestauranter* (*luxurious lunch restaurants*) as *déjeuner exclusive restaurants*, literally meaning “lunch luxurious (fem. sing.) restaurants” (masc. sing.). Bing also has serious problems dealing with the complex Norwegian compound noun phrase, *8 kilometer lange flotte rullesteinsstrender* (*lovely 8-kilometer-long pebble stone beaches*), putting the translated words into a seemingly random order: *8 km plages de galets long magnifique* (*8 km beaches of pebbles long lovely*).

Likewise, Google turns the Norwegian noun phrase, *sjarmerende fargerike murhus* (*charming colourful brick houses*) into *charmantes maisons colorées de briques*, i.e. the adjective *colorées* is wrongly positioned between the noun *maisons* (*houses*) and the prepositional phrase *de briques* (*of bricks*), which in this case are tightly linked and should not be separated.

Longer clauses with heavily modified subjects and complex verb phrases yield incoherent and almost incomprehensible translations with both programmes.

Problems with word order and sentence structure are probably the most disturbing error type for a reader who is trying to grasp the text. They are also the hardest to describe, as it takes some knowledge of French syntax to be able to explain them. Therefore, I expected the students’ comments to be more precise and concrete on the morphological errors than on the word order errors.

Procedures

Task Design and Data Collection

The task offered to the students consisted of reading two machine-translated versions of the same text and deciding, in groups, which translated version was best while explaining, in as much detail as possible, why one was better than the other. The task resembled traditional error correction, which has been used in foreign language learning for a long time and is often based on texts containing errors deliberately added by the teacher. When working with text produced by MT programmes, however, students work with errors and shortcomings that they encounter on a daily basis and need to be aware of. Moreover, in the present activity, students were asked to detect errors but not specifically to correct them (i.e. students were not asked to post-edit the machine-translated texts). Instead, machine-translated output was used to prompt students to talk about language.

The teacher had informed the students about the project and its purpose in a prior lesson. She had also instructed them to read the Norwegian source text to become more familiar with the content when it was presented with the two translated versions. In class, during the project work, the students were given about 25 minutes to read the two translated versions individually and

to jot down errors and shortcomings that they found while reading. Then, they were divided into seven groups of 2–4 students each and given another 20 minutes for group discussion and comparison of the individual findings. The students were informed that they could use Norwegian (L1) in their discussions. In order to facilitate the discussion and subsequent data analysis, the lines in the Google-translated text were numbered from 1 to 25, while the lines in the Bing-translated text were labelled A–Y. A sheet with task instructions and guiding questions that specified elements the students could look for (see Appendix) was distributed to all students before the individual reading, and they were prompted to use the sheet during group work. I was present the entire time and could answer any questions the students had about the task or the questions on the sheet. The teacher ensured that the participants were homogeneous within each group in terms of their competence level in French. The students did not know that they were divided into groups on the basis of achievement level. Table 5.1 provides an overview of the composition of the groups.

The group conversations were recorded using mini-recorders placed on each group's table. The data thus consisted of seven recorded group conversations of approximately 20 minutes each.

The teacher was instructed to act normally, i.e. to intervene with the groups when relevant or necessary and ask them any questions she considered useful for their work with the task. Apart from this, the teacher did not receive any instructions on what to do or say, nor did we discuss in any detail how the students could be guided in that particular activity.

Data Analysis

Once the data had been collected, I listened to the material with the research questions in mind, and all instances of collaborative talk, operationalised in this study as language-related episodes (LREs) (see Swain and Lapkin 2001, 104) were transcribed. Whereas Swain and Lapkin's concept of LREs refers to episodes in which students talk about their own or their peers' linguistic output, I used it in this study to refer to any part of a conversation where the students talked about and reflected upon the MT linguistic

Table 5.1 Group composition.

<i>Group</i>	<i>Achievement level in French</i>	<i>Number of students</i>
A	High	2 (both girls)
B	Medium	3 (two girls, one boy)
C	Medium	4 (two girls, two boys)
D	Medium	3 (all girls)
E	Low	4 (all boys)
F	Low	3 (all girls)
G	Low	4 (one girl, three boys)

output. The underlying principle, however, was the same: LREs refer to episodes in which students discuss or comment on language use and linguistic choices; in this case, in regard to machine translations. Time stamps for the beginning and end of these episodes helped determine the time each group spent on collaborative talk.

In order to answer Research Question 1 (how students talk about language and the extent to which they use grammatical terms), I first read through the transcripts to obtain a general impression of how the groups talked about the texts. I then reread the transcripts while identifying and counting all instances of grammatical terms for each of the groups. Grammatical terms included references to parts of speech, grammatical gender, verb conjugation, concord and so on. All instances were counted, regardless of whether or not the terms were used correctly.

In order to answer Research Question 2 (what kinds of errors the students detected), the transcripts were coded in terms of whether they focused on morphological, lexical, syntactic or mechanical linguistic errors or shortcomings. These categories were largely pre-defined, based on the error types present in the text (see above) and on previous LRE taxonomies (Fortune 2013, 173; Gutiérrez 2013; Storch 2008). A few subcategories that were not identified in the pre-analysis of the texts emerged from the data, including comments about numerals (lexical), punctuation (mechanical) and unspecified comments. The latter category comprised episodes in which the students mentioned that something was wrong, but they were not specific enough about the error for their statements to be placed in a specific error-type category. One example of this is when a student stated that “this sentence is kind of strange”. Episodes in which the students merely discussed the text without referring to (presumed) errors or shortcomings were counted as LREs, and hence as collaborative talk, but classified simply as discussion about text. Examples in this category primarily include instances where the students asked each other about the meanings of words; for example, when one student said, “I don’t know what ‘cadre’ means”, and another (wrongly) responded, “I think it is ‘part’”.

The findings for each group were compared in order to inform Research Question 3, which addressed differences between achievement levels. The number of grammar terms was compared across groups, and comments for each error type were also compared to identify which errors were discussed by most of the groups and which were only detected by specific groups.

Results

Table 5.2 gives a list of all instances of grammatical terminology used by each group. Table 5.3 illustrates the time spent on collaborative talk and the number of unique LREs identified for each group and provides an overview of the error types detected and commented on by each group.³ Most of the time not spent on collaborative talk was devoted to discussing off-task

Table 5.2 Grammatical terminology used by each group.

<i>Group</i>	<i>Terms used (number of occurrences in parentheses)</i>
A (high)	adjective (4), amalgam (1), article (4), conjugate/conjugation (4), feminine ending (1), feminine noun (1), feminine plural (1), indefinite (1), noun (4), phrase / fixed expression (1), plural (2), preposition (3), present tense (1), subordinate clause (1), verb (5)
B (medium)	adjective (1), article (4), definite (1)
C (medium)	adjective (2), conjugate (5), number (1), person (1), subject (1), verb (2)
D (medium)	adjective (1), verb (1), verb conjugation (1)
E (low)	conjugation (1), verb (1)
F (low)	conjugation (1), feminine (1), masculine (1), plural (1), verb (1)
G (low)	–

matters (e.g. what to do during the weekend). The time spent on collaborative talk therefore represents the groups' time-on-task fairly well, with a low number of minutes (see Table 5.3) indicating that a group did not work for very long on the task. Notably, there were some episodes of silence, which might represent time spent on-task since the students may have been working individually: silently reading the text or simply thinking. Below I will present and discuss the observations made from the data and link them to each of the research questions.

Research Question 1: How Do the students Talk About the Texts, and to What Extent Do They Use Grammatical Terms?

The most salient finding was that, in most groups, students' comments about the text were very general. They often made claims without justifying them with concrete examples from the text. The following excerpt⁴ from group C is illustrative in this respect:

S1: I feel that Google in a way has OK words compared with Bing

S2: Yes, Bing tries to write adult French

S3: Yes, like advanced French

[...]

S1: Bing varies the words more

S2: But Google has more appropriate words, kind of

Group C (medium)

It is difficult to determine whether this lack of concrete examples was caused by the students not having paid close enough attention to the task instructions or by the fact that they were not able to find illustrative examples to support their impressions. One aspect points to the latter explanation,

Table 5.3 Time spent on collaborative talk and error types discussed for each student group.

<i>Group (achievement level)</i>	<i>Time spent on collaborative talk</i>	<i>Main error types discussed</i>	<i>Detailed error types discussed</i>	<i>No. of LREs per (assumed) error type*</i>	<i>No. of other LREs**</i>
A (high)	29 min 30 sec	Morphological	Agreement errors	4	20
			Omission (e.g. of articles)	4	
			Verb conjugation	1	
			Preposition use	1	
			Article use	1	
		Lexical	Lexical choice—proper names	2	
			Lexical choice—other	3	
			Untranslated words	3	
			English words	1	
			Use of numerals	2	
		Syntactic	Adjective placement	4	
			Word order in general	2	
			Punctuation	1	
B (medium)	14 min 30sec	Mechanical			6
		Morphological Lexical	Omission (e.g. of articles)	2	
			Lexical choice—proper names	2	
			Lexical choice—other	5	
			Untranslated words	2	
			English words	1	
			Use of numerals	1	
			Adjective placement	1	
			Verb conjugation	1	
			Lexical choice—proper names	3	
			Lexical choice—other	1	
			Untranslated words	1	
C (medium)	20 min 30 sec	Syntactic Morphological Lexical	English words	1	9
			Use of numerals	1	
			Adjective placement	1	
			Verb conjugation	1	
			Lexical choice—proper names	3	
			Lexical choice—other	1	
			Untranslated words	1	
			English words	1	
			Use of numerals	2	
			Adjective placement	1	
		Syntactic Mechanical	Punctuation	2	

(Continued)

Table 5.3 (Continued)

<i>Group (achievement level)</i>	<i>Time spent on collaborative talk</i>	<i>Main error types discussed</i>	<i>Detailed error types discussed</i>	<i>No. of LREs per (assumed) error type*</i>	<i>No. of other LREs**</i>
D (medium)	9 min 00 sec	Lexical	Lexical choice—proper names	3	7
			English words	1	
			Use of numerals	1	
			Adjective placement	1	
			Punctuation	1	
E (low)	9 min 30 sec	Syntactic Mechanical Lexical	Lexical choice—proper names	2	4
			Lexical choice—other	1	
			English words	1	
			Use of numerals	1	
			Adjective placement	1	
F (low)	19 min 30sec	Syntactic Mechanical Morphological Lexical	Punctuation	1	7
			Agreement	3	
			Preposition use	2	
			Lexical choice—proper names	3	
			Lexical choice—other	1	
G (low)	13 min 00 sec	Syntactic Lexical	Use of numerals	1	5
			Untranslated words	1	
			English words	1	
			Word order in general	1	
			Lexical choice—proper names	2	
			Lexical choice—other	1	
			Use of numerals	1	

* Repeated episodes (i.e. episodes where the participants came back to the same example as the one discussed before) were counted only once. Thus, when group A discussed agreement errors in four LREs, they discussed four different instances of agreement errors.

** Includes *unspecified comments* and *discussion about text* (see the data analysis section)

however: When the teacher occasionally asked the students to justify their claims by giving illustrative examples, the students were unable to do so:

T: Can you give some examples?

S: Ah . . . no . . .

Group F (low)

Grammatical terminology occurred infrequently and sporadically, except in the high-achieving group (see Table 5.2). This group used grammatical terms in a nearly systematic way. In their 30-minute talk, they used 15 different grammar terms to comment on and discuss phenomena such as noun-adjective concord, omission of articles and verbs, verb conjugation and the position of adjectives. The following excerpts illustrate their talk about noun—adjective concord and article omission, respectively:

S1: What I also noticed, was that—it is not so important—but when it says that one of the cities was most visited, it says visités

S2: But ville, isn't that a feminine noun?

S1: Yes, so it should have been feminine plural

S1: Yes, and there should have been an article before aristocrates

S2: Should have been des, then, if it is indefinite

Group A (high)

Even the high-achieving group was occasionally hesitant about terminology. For instance, one of the girls first correctly talked about “articles”. The other girl objected, claiming that “it is more a preposition”, and in the subsequent dialogue the term “articles” was erroneously replaced by “prepositions”.

The other groups rarely made use of grammatical terminology. The number of occurrences ranged from zero in group G to 12 in group C. When they did use such terminology, it was often in a vague or even erroneous manner. For example, in one of the medium-achieving groups, one student said: “In French, *belle* and *petite* and *grand* and so on can occur in front of a verb, can't they?” The student is obviously referring to the rule whereby some French adjectives (including the ones mentioned) are placed in front of the noun. He did not use the term “adjectives” but rather mentioned some examples, and he confused the terms “noun” and “verb”.

As previously observed by Hansejordnet (2009), students' lack of linguistic terminology and grammatical understanding sometimes seemed to hamper their analysis of the translated text. For example, in the following excerpt, the students could see that there was something wrong with the verb *regarder* (English *look*, Norwegian *se, titte*), which occurs in the infinitive in Google's version. They related this to the verb form in the Norwegian

source text (lines 5–6), but they did not identify it as an imperative, which could have helped them in the analysis of the French text.⁵

- 1) S1: They don't conjugate all the verbs, either
- 2) S2: Where exactly?
- 3) S1: Line 18
- 4) S3: But here, in Bing, they have conjugated it and it should . . .
[. . .]
- 5) S2: If you look at the original it is not conjugated
- 6) S3: Yes, it is, it says titt,⁶ doesn't it. It isn't à tittle,⁷ kind of.
- 7) S2: I think that's correct, because it's not conjugated according to subject
- 8) S3: So the other is wrong?
- 9) S2: One of them has got to be wrong
- 10) S3: The inflections are completely different

Group C (medium)

This example also illustrates another observation: When the students spoke about the translated texts, they often spoke as if one programme was necessarily right and the other wrong. Similarly, when they talked about rules, they did not seem to realise that there could be exceptions, or that there could be instances in which the rule did not apply. It is not clear why such misconceptions arise, but one possible explanation could be that students are largely exposed to simplified input constructed to illustrate rules and are rarely confronted with authentic language.

Research Question 2: What Kinds of Errors Do They Talk About?

Lexical errors received by far the most attention from the students. All the student groups commented, on their own initiative, on Google Translate's choice of heading, where Nice was translated to *agréable*. Nevertheless, four groups ignored the fact that Bing actually made the same kind of mistake, but in the body text, and by choosing *belle* instead of *agréable*. Three groups at different achievement levels (A, D and F) noticed this:

- S1: And in the second paragraph in Bing, they have translated Nice to belle
 S2: Where?
 S1: Line E
 S2: Yes!

Group D, medium-achieving

In addition, the high-achieving group immediately captured the importance of context for choosing the correct translation. They stated that "There's nothing that could indicate that it's the name of a place" and that "there

is no context, in a way". They also stressed that Nice was not a Norwegian word and that the programme "didn't make that mistake in the rest of the text, because then they had the context". These students thus "excused" the translation programme for the erroneous title because it did not have the context needed to render the source text appropriately.

Five groups also commented on the translations of "rivieraen" (*the Riviera*). Google kept the word *Riviera*, while Bing used the correct French term *le Côte d'Azur*. The students were obviously familiar with this term, recognising that "that's what it's actually called" (group D, medium).

Another issue that all the groups commented on was the use of numerals by the two translation programmes, and more precisely how they express centuries. The students were surprisingly concerned with numerals, perhaps because they were concrete and easy to look for. They noticed that while Bing chose Roman numerals, Google used regular numbers, yet they disagreed on which solution was best. Some called Bing's use of Roman numerals "fancy", while others found it "silly".

Four groups commented on the untranslated terms, and six commented on the use of the English word *must*, which also occurred in the original Norwegian text. While some considered this erroneous, others suggested that the English word *must* might be used as a loan word in French just as it is in Norwegian.

Most comments on the lexical level addressed proper names or untranslated words. There were some exceptions, however. For example, the high-achieving group expressed some doubts about the word *occupée* (these doubts are justified; see above), and one of the medium-achieving groups expressed some doubts about the word *animée* (these doubts are not justified, as the machine's choice of word was appropriate). Overall, there were very few reflections on different word meanings. A common misconception about languages is that one word has only one meaning. In the following excerpt, this is clearly the conception guiding the students' talk about the word *même*. They are not only unfamiliar with the different meanings of *même*⁸ they also do not seem to realise that this word may have several meanings.

S1: What's "même"?

S2: I guess that's like "as"

S1: Yes, (SCEPTICAL) because here they have used it as "even". Even a small pharmacy

S3: There is something called "plus même". That's "in addition".

S1: It's half way right in a way, if they lack "plus"

Group B (medium)

The students here showed a lack of awareness that words can be polysemous: After having established the fact that *même* means *samellike*, they seemed to assume that it could not have other meanings, such as *even*.

The students rarely discussed nuances in lexical choices. One of the medium-achieving group's discussions on the adjectives *petites* and *minuscules* constitutes a notable exception:

- S1: But here in Bing they have written "petites boutiques" and here they have written "minuscules boutiques"
 S2: "Minuscules" is also possible
 S1 & S3: Is it?
 S2: "Minuscules", yes (PRONOUNCES IT IN ENGLISH) Yeah, that's possible! That's a word
 S1: But it's supposed to be tiny shops
 S2: Yes, "minuscule" means "tiny"
 S1: Alright, then it's Google [which is right]
 S2: [No, I think both to be honest]
 S1: You can also say "petites"
 S3 (TO S2): Do you actually know what you're talking about?
 S1: I would have written "petites"
 S3: I would also have written "petites". [But it may be because we have not yet learnt . . .]
 S2: [That's because you don't have as broad a vocabulary as I have!]
 S1: I suppose it's because we haven't come that far

Group B (medium)

In this excerpt, S1 at first seems to share the quite commonly held belief that there is a one-to-one correspondence between words in different languages. She claims that *minuscules* is probably wrong because it's supposed to be 'tiny' and the word she knows for 'tiny' is *petit*. S2 challenges this view by claiming that *minuscules* is also right, as he knows the meaning (he compares it with the English word *minuscule*, which he knows). S1 and S3 then conclude that both words may be right, and that *minuscules* is just unfamiliar to them because they have not yet learnt it.

We see sporadic examples of this kind of metalinguistic awareness among other students as well. The low-achieving group discussed the programmes' translations of "the old town in Nice"—"*la vieille ville de Nice*" and "*vieux Nice*", respectively—and concluded that both may be right because "there are several ways of saying things" (group G, low-achieving).

All groups except for group G also tried to comment on grammatical errors, but such instances were very limited in number (except for the high-achieving group), and most comments were general and not linked to concrete examples. Most of the concrete comments concerned verb conjugation and adjective placement. The students' observations were sometimes appropriate, while at other times they were vague and/or inaccurate. In the following example, the students seem to ignore the relevance of co-text, as they do not realise that the two programmes have chosen different nouns, and therefore the ending on the adjective *somptueux* also differs:

- S1: And here in line 3, we have “somptrueuse” and “somptrueux”, which one is correct?
S2: That’s unclear
S3: It’s feminine and masculine, isn’t it?
S1: Right, because when you add the ‘e’ . . .
S2: Bing says feminine. That’s wrong, we say!

And in the following excerpt, the students correctly assume that the adjective *français* should be placed after the noun, but they do not manage to explain the reason for this:

- S1 (REFERRING TO THE ADJECTIVE *FRANÇAIS*): That “French” should come after, shouldn’t it? It’s an adjective
S2: Yes . . .
S1: But here it’s the other way around. Here it’s after and there it’s before
S2: Where exactly?
S1: Line 13
S3: Where?
S1: Where that Chablis-thing is
S3: I feel that Google is correct. I believe it’s right that ‘French’ should come after
S1: So do I
S3: I think Google is right in this case, because. . . Chablis, I think you say Chablis in French and English and everything

Although there were, in most groups, a few examples of morphological error detection, this was clearly not the students’ favourite topic. For instance, when a student in group D attempted to start a discussion on grammar points, the others refused:

- S1: What about like verb inflection and that kind of stuff?
S2: I don’t know. I don’t know anything about that.

Group D (medium)

Episodes like this, where some students tried to bring up linguistic points they had noticed but the others did not respond, occurred in several of the groups, demonstrating that the mentioning of morphological aspects tended to elicit only limited engagement.

There were some comments on word order and sentence structure, but these were often vague. Still, the students recognised that the jumbled word order was the main problem with the MT texts:

- S1 (READING FROM THE SHEET): What errors and shortcomings are most detrimental for the understanding of the text?
S2: The word order!

Group F (low)

They also realised that the longer the clause, the worse the translation:

- S1: Both are better with single words. Entire texts just yield lots of mistakes
 S2: Yeah, the longer the sentence, the worse it gets.

Group C (medium)

This shows that even though students have difficulties explaining the problems with word order, they demonstrate critical awareness when it comes to what the MT programme can or cannot do.

In sum, the majority of the groups spent considerable time discussing lexical errors, mostly obvious errors such as Google's heading and untranslated words. They also spent much time discussing numerals, and most also discussed the translation of Riviera. In addition, there were some sporadic comments from individual students about grammar issues. These comments were sometimes vague and/or erroneous and sometimes clear, but they were rarely picked up by the rest of the group.

Research Question 3: What Differences Are There between the Different Achievement Levels?

The most striking difference between the achievement levels was that the high-achieving group commented on grammar and used technical metalanguage, while the others in general did not. These two aspects (grammar comments and technical metalanguage) are linked: Gutiérrez (2013) found that students can easily talk about lexical aspects without resorting to technical terms, but that they need them while discussing morphological and syntactic aspects (Gutiérrez 2013, 163). Also, as the examples provided above have shown, the high-achieving group was far more concrete in their discussion of the texts, supporting their claims with examples from the texts. In all the other groups, general claims such as "Bing is better, because they have more coherence" dominated. The students did not make any attempts to support their claims with examples from the texts, even though the instructions for the task asked them to be as detailed as possible.

Another difference between the groups was the amount of time spent on collaborative talk, ranging from approximately nine minutes in groups D and E to nearly 30 minutes in the high-achieving group (see Table 5.3). The high-achieving group continued working even after all the other groups had stopped. Many of the other groups spent a great deal of time joking or discussing off-task topics. However, this difference does not necessarily have anything to do with achievement level. The low-achieving group G took the task seriously, and although they spent quite some time off-task, this time did not consist of joking but rather of silence, as the students felt lost and did not know what to say. In these cases, they eventually called on the teacher for help and thus moved back to time-on-task. It seems that factors other than achievement level influenced the time spent on the task. While

the high-achieving group probably could have benefited from more time, the other groups would definitely not have. Nonetheless, all the groups would have benefited from more guidance (see next section).

What is probably more linked to achievement level is how the students perceived their own competence. All the student groups expressed insecurity and lack of knowledge (“I don’t know, but . . . this is hard for us to determine . . .”), which is quite normal in the situation, but the low-achieving group G displayed extremely low self-confidence when it came to French, and they had a very negative perception of their own competence. In this group, one student started the conversation by stating, “I don’t know anything. I can’t tell the difference between the two texts.” And when they came to the question, “Which errors are most detrimental to the intelligibility of the text?” they said, “Ha ha, the fact that it’s French!” Although the purpose of the task was given in the beginning of the class, this group did not understand why they had been given it: “I don’t know any French, why do they ask us to correct French texts?”

Discussion and Conclusion

This study set out to examine how L3 students talk about language when such talk is triggered by machine-translated texts. The idea was to explore whether MT texts could engage students in metalinguistic talk and help them adopt an analytical and critical perspective on language use. The students’ discussions, which took place in small achievement-based groups, were recorded and subsequently analysed with regard to the students’ use of grammatical terms, comments about various error types and time spent in collaborative dialogue.

Some limitations of the study need to be mentioned. First, the study was conducted only once in a single class, so caution must be exercised when interpreting the results. Second, the task was presented and explained to the students, but it was not modelled. A modelling phase would have helped ensure that the students knew what to do and how to do it, and uncertainty about the reason why students provided few concrete examples (lack of linguistic knowledge or simply lack of attention to the task instructions) could have been reduced. Third, the teacher did not receive any training in how to manage this type of task prior to the observed lesson. The full potential of the task was therefore not realised. Yet another important issue is that this project focused on group work only. Individual notes were not kept, but such notes could have yielded valuable additional information regarding what individual students noticed compared to what aspects were discussed in groups.

The findings indicate that most student groups talk about the texts and the errors in them using general vocabulary and without referring to concrete examples. While the students are most concerned with lexical errors, their comments are mostly restricted to proper names and untranslated

words. They rarely discuss nuances in semantic meaning. Regarding the differences between achievement levels, the most salient findings are that the high achievers spend considerably more time in collaborative dialogue than those in the other groups, and they use notably more grammatical terminology. The high-achieving group clearly find the task useful since they want to continue working after the time elapses. This mirrors Storch's (2008) findings: The level of engagement is essential for a fruitful outcome, but of course the level of engagement is influenced by the students' achievement level. Considering the level of engagement and the time spent in successful collaborative dialogue, it is reasonable to assume that the high achievers learn more from the task than those in the other groups.

The analysis further reveals that many of the students have simplified ideas about language. They tend to believe that there is only one correct answer, ignore the relevance of co-text and believe there is a one-to-one relationship between words in different languages and between a word and its meaning. Such misconceptions are found in the groups with low and medium achievement levels but not in the high-achieving group. Nevertheless, in the data, we can see examples of learning or consolidating learners' previous knowledge in groups of all levels: In one of the low-achieving groups, the students acknowledge that several solutions can be correct ("there are several ways of expressing things"); in the high-achieving group, the relevance of context (*agréable*) is understood; and in one medium-achieving group, the conception of a one-to-one relationship between words across languages is challenged by other group members (*petit- minuscule*). This is a good example of the co-construction of knowledge that can occur in collaborative dialogue. The discussion of lexical meanings also led the group to acknowledge the value of having a broad vocabulary.

The task could potentially have triggered the learning of several other linguistic aspects had the students received more scaffolding and guidance. They could have, for instance, learned more about polysemy (e.g. *même*) and differences between Norwegian and French word order and sentence structuring.

Nevertheless, these observations suggest that machine-translated texts can potentially be used to both trigger and develop more nuanced and appropriate conceptions about language, also with beginners. In this respect, the findings of the present study are in line with those in Ninõ's (2008) study involving advanced learners, suggesting that working with MT texts has a beneficial effect for this group. In addition, the present study demonstrates how MT texts can be used to help learners learn more about the language and not just to communicate better (see Garcia and Pena 2011). However, the present study also shows that it is essential to reflect upon how machine-translated texts can best be introduced with this aim. Clearly, it is not sufficient to introduce the tool and assign students related tasks to be completed. There is a clear need for scaffolding techniques and teacher guidance, even

when students work in groups. To a certain extent, students can guide each other, but the data show that in cases where none of the students are really sure about the point they are discussing, group conversation tends to reinforce misconceptions about language, as we saw in the discussion about the meaning of “*même*” and in the high-achieving group’s confusion between articles and prepositions. Moreover, elements posed by one student are not always picked up by the others—in fact, they are often ignored. This happens most often when a student makes a comment about morphological aspects. Unlike in Storch’s (2008) study, the mentioning of morphological aspects thus tends to elicit only limited engagement. These contrasting findings can probably be explained by the ages and interests of the learners; while Storch’s participants were university students who had chosen to study a foreign language, the participants in the present study were secondary school students for whom foreign language is a compulsory subject. The majority are probably not particularly interested in grammar.

Thus, the outcome of the collaborative dialogue is not always successful, and the role of the teacher therefore seems to be of utmost importance. However, given the general teacher reluctance towards the use of these tools, the novelty of bringing them into the classroom for learning purposes and the general insecurity teachers feel about teaching metalinguistic awareness (Vold 2017), it seems obvious that teachers would also need training in how to guide the students in this type of work. Tasks like the one presented in this chapter should therefore be used in the teacher training process. Such tasks have the potential to raise teacher trainees’ awareness of the importance of metalanguage as well as cross-linguistic similarities. These types of activities are greatly needed in teacher training, as scholars such as Otwinowska (2014) and E. Ellis (2012) have pointed out. What seems equally clear is that once a teacher has been trained in how to organise these kinds of tasks, the approach must be systematic and repeated. The students may not learn as much from one first, single try-out, especially if this try-out is not followed up by a consolidation phase where students sum up and elaborate on points they have learned. Going back to our data, such points could be, in our case, (a) the relevance of context, (b) the importance of having a broad vocabulary, (c) the existence of multiple but equally appropriate solutions, (d) polysemy and (e) differences between French and Norwegian word order and sentence structuring.

In sum, it seems that the use of MT texts to develop students’ metalinguistic awareness has considerable potential, but training, scaffolding techniques and guidance from the teacher are of paramount importance. Systematic repetition of the same type of task, followed each time by a consolidation phase, is also crucial. With such guidance, the discussion and evaluation of MT texts could foster metalinguistic awareness. To more precisely determine the value of this kind of task, further research focusing on long-lasting interventions that include modelling, scaffolding and a

consolidation phase is required. The need for such action-based research is considerable, as MT is definitely here to stay and students use and will undoubtedly continue to use it. In this respect, it is vital that not only students' metalinguistic awareness be addressed but also their critical awareness of technology.

Notes

- 1 These two terms are often used interchangeably. In this chapter, I use metalinguistic awareness in order to stress the idea that this type of knowledge is explicit, i.e. something the learner is conscious about and aware of.
- 2 The letters in parentheses behind each example indicate that the example is taken from either Google Translate (GT) or Bing Translator (BT).
- 3 Repeated LREs, in which the learners come back to an observation discussed before, were counted only once.
- 4 Underlining indicates word stress. [Brackets] indicate start and end points of overlapping speech. Capital letters annotate non-verbal activity. T = Teacher, S1 = Student 1, S2 = Student 2, etc. The excerpts are translated from Norwegian by the author.
- 5 In French, the imperative can be expressed by an imperative form of the verb (*regardez*, as chosen by Bing) or by the infinitive (*regarder*, as chosen by Google). In some contexts, they are interchangeable; in other contexts one of them would be the preferred or only choice. The students are correct in their observation that, in this case, Google's *regarder* does not work well (see Appendix).
- 6 Imperative form of *to peek/look*
- 7 Infinitive, *to peek/look*
- 8 The word *même* has several meanings in French, of which *same/like/as* and *even* are two of the most frequent.

References

- Cenoz, Jasone, Britta Hufeisen, and Ulrike Jessner. 2001. "Towards Trilingual Education." *International Journal of Bilingual Education and Bilingualism* 4 (1): 1–10. <https://doi.org/10.1080/13670050108667714>
- Clifford, Joan, Lisa Merschel, and Joan Munné. 2013. "Surveying the Landscape: What Is the Role of Machine Translation in Language Learning?" *Revista d'innovació Educativa* 10: 108–21. <https://doi.org/10.7203/attic.10.2228>
- De Angelis, Gessica. 2011. "Teachers' Beliefs About the Role of Prior Language Knowledge in Learning and How These Influence Teaching Practices." *International Journal of Multilingualism* 8 (3): 216–34. <https://doi.org/10.1080/14790718.2011.560669>
- Ellis, Elizabeth. 2004. "The Invisible Multilingual Teacher: The Contribution of Language Background to Australian ESL Teachers' Professional Knowledge and Beliefs." *International Journal of Multilingualism* 1 (2). <https://doi.org/10.1080/14790710408668181>
- . 2012. "Language Awareness and Its relevance for TESOL." *University of Sydney Papers in TESOL* 7, 1–23.
- Ellis, Rod. 2004. "The Definition and Measurement of L2 Explicit Knowledge." *Language Learning* 54: 227–75. <https://doi.org/10.1111/j.1467-9922.2004.00255>
- Fortune, Alan. 2013. "The Role of Metalanguage in the Performance of a Sequence of Collaborative Output Tasks by Five L2 Learner Dyads." In *The Metalinguistic*

- Dimension in Instructed Second Language Learning*, edited by Karen Roehr and Gabriella A. Gánem-Gutiérrez, 171–95. London: Bloomsbury.
- Garcia, Ignacio, and María I. Pena. 2011. “Machine Translation-assisted Language Learning: Writing for Beginners.” *Computer Assisted Language Learning (CALL)* 24 (5): 471–87. <https://doi.org/10.1080/09588221.2011.582687>
- Gaspari, Federico, Hala Almaghout, and Stephen Doherty. 2015. “A Survey of Machine Translation Competences: Insights for Translation Technology Educators and Practitioners.” *Perspectives: Studies in Translatology* 23 (3): 333–58. <https://doi.org/10.1080/0907676X.2014.979842>
- Golonka, Ewa, Anita Bowles, Victor Frank, Dorna Richardson, and Suzanne Freynik. 2014. “Technologies for Foreign Language Learning: A Review of Technology Types and Their Effectiveness.” *Computer Assisted Language Learning (CALL)* 27 (1): 70–105. <https://doi.org/10.1080/09588221.2012.700315>
- Gutiérrez, Xavier. 2013. “Metalinguistic Knowledge in Language-related Episodes with Covert Metalinguistic Activity.” In *The Metalinguistic Dimension in Instructed Second Language Learning*, edited by Karen Roehr and Gabriela A. Gánem-Gutiérrez, 147–71. London: Bloomsbury.
- . 2016. “Analyzed Knowledge, Metalanguage, and Second Language Proficiency.” *System* 60: 42–54. <https://doi.org/10.1016/j.system.2016.06.003>
- Hansejordet, Ingunn. 2009. *Grammatikken i spanskfaget* [Grammar in the Subject of Spanish]. Fokus på språk 14. Halden: Fremmedspråksenteret.
- Hasselgård, Hilde. 2018. “Language Contrasts, Language Learners, and Metacognition: Focus on Norwegian Advanced Learners of English.” In *Metacognition in Language Learning and Teaching*, edited by Åsta Haukås, Camilla Bjørke, and Magne Dypedahl. New York: Routledge.
- Haukås, Åsta. 2012. “Lærarhaldningar til språklæringsstrategiar [Teachers’ Beliefs About Language Learning Strategies].” *Norsk Pedagogisk Tidsskrift* 2: 115–29.
- . 2016. “Teachers’ Beliefs About Multilingualism and a Multilingual Pedagogical Approach.” *International Journal of Multilingualism* 13 (1): 1–18. <https://doi.org/10.1080/14790718.2015.1041960>
- . 2018. “Metacognition in Language Learning and Teaching: An Overview.” In *Metacognition in Language Learning and Teaching*, edited by Åsta Haukås, Camilla Bjørke, and Magne Dypedahl. New York: Routledge.
- Heimark, Gunn Elin. 2013. “Praktisk tilnærming i teori og praksis. Ungdomsskolelæreres forståelse av en praktisk tilnærming i fremmedspråkundervisningen [An Action-Oriented Approach in Theory and Practice. Lower Secondary School Teachers’ Perception of an Action-Oriented Approach in Foreign Language Teaching].” PhD diss., University of Oslo, Oslo.
- Jessner, Ulrike. 2008. “A DST Model of Multilingualism and the Role of Metalinguistic Awareness.” *The Modern Language Journal* 92 (2): 270–83. <https://doi.org/10.1111/j.1540-4781.2008.00718>
- . 2018. “Metacognition in Multilingual Learning: A DMM Perspective.” In *Metacognition in Language Learning and Teaching*, edited by Åsta Haukås, Camilla Bjørke, and Magne Dypedahl. New York: Routledge.
- Jolley, Jason R., and Luciane Maimone. 2015. “Free Online Machine Translation: Use and Perceptions By Spanish Students and Instructors.” In *Learn Languages, Explore Cultures, Transform Lives: Selected Papers from the 2015 Central States Conference on the Teaching of Foreign Languages*, edited by Aleidine J. Moeller, 181–200. Eau Claire, WI: Robert M. Terry.

- Lyle, Susan. 2008. "Learners' Collaborative Talk." In *Encyclopedia of Language and Education*. 2nd ed., Vol. 3, edited by Nancy Hornberger, 1022–33. Boston: Springer.
- Niño, Ana. 2008. "Evaluating the Use of Machine Translation Post-editing in the Foreign Language Class." *Computer Assisted Language Learning (CALL)* 21 (1): 29–49. <https://doi.org/10.1080/09588220701865482>
- . 2009. "Machine Translation in Foreign Language Learning: Language Learners' and Tutors' Perceptions of Its Advantages and Disadvantages." *ReCALL* 21 (2): 241–58. <https://doi.org/10.1017/S0958344009000172>
- Norwegian Directorate of Education and Training. 2006. "Subject Curriculum for Foreign Languages." Norwegian Directorate of Education and Training. Accessed June 29, 2016. www.udir.no/kl06/FSP1-01/Hele/Kompetansemal/foreign-language-level-ii?lplang=eng
- Otwinowska, Agnieszka. 2014. "Does Multilingualism Influence Plurilingual Awareness of Polish Teachers of English?" *International Journal of Multilingualism* 11 (1): 97–119. <https://doi.org/10.1080/14790718.2013.820730>
- Roehr, Karen. 2008. "Metalinguistic knowledge and language ability in university-level L2 learners." *Applied Linguistics* 29 (2): 173–99. <https://doi.org/10.1093/applin/amm037>
- Roehr, Karen, and Gabriela A. Gánem-Gutiérrez. 2013. *The Metalinguistic Dimension in Instructed Second Language Learning*. London: Bloomsbury.
- Somers, Harold. 2001. "Three Perspectives on MT in the Classroom." In *MT Summit VIII Workshop on Teaching Machine Translation*, Santiago de Compostela, 25–29.
- Somers, Harold, Federico Gaspari, and Ana Niño. 2006. "Detecting Inappropriate Use of Free Online Machine-translation By Language Students—A Special Case of Plagiarism Detection." In *Proceedings of the 11th Annual Conference of the European Association for Machine Translation*, Oslo, Norway, 19–20 June 2006, 41–8.
- Steding, Søren. 2009. "Machine Translation in the German Classroom: Detection, Reaction, Prevention." *Die Unterrichtspraxis/Teaching German* 42: 178–89. <https://doi.org/10.1111/j.1756-1221.2009.00052>
- Storch, Neomy. 2008. "Metatalk in a Pair Work Activity: Level of Engagement and Implications for Language Development." *Language Awareness* 17 (2): 95–114. <https://doi.org/10.1080/09658410802146644>
- Swain, Merrill. 1998. "Focus on Form Through Conscious Reflection." In *Focus on Form in Classroom Second Language Acquisition*, edited by Catherine Doughty and Jessica Williams, 64–82. Cambridge: Cambridge University Press.
- . 2000. "The Output Hypothesis and Beyond: Mediating Acquisition Through Collaborative Dialogue." In *Sociocultural Theory and Second Language Learning*, edited by James Lantolf, 97–114. Oxford: Oxford University Press.
- Swain, Merrill, Penny Kinnear, and Linda Steinman. 2015. *Sociocultural Theory in Second Language Education*. 2nd ed. Bristol: Multilingual Matters.
- Swain, Merrill, and Sharon Lapkin. 2001. "Focus on Form Through Collaborative Dialogue: Exploring Task Effects." In *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*, edited by Martin Bygate, Peter Skehan, and Merrill Swain, 99–118. London: Longman.
- Thomas, Jacqueline. 1988. "The Role Played By Metalinguistic Awareness in Second and Third Language Learning." *Journal of Multilingual & Multicultural Development* 9 (3): 235–46. <https://doi.org/10.1080/01434632.1988.9994334>

- Vold, Eva Thue. 2017. "Qualifying L3 Teachers: Is Teacher Training Enough?" *International Journal of Educational Research* 82: 40–53. <https://doi.org/10.1016/j.ijer.2016.12.002>
- Williams, Lawrence. 2006. "Web-Based Machine Translation as a Tool for Promoting Electronic Literacy and Language Awareness." *Foreign Language Annals* 39 (4): 565–78. <https://doi.org/10.1111/j.1944-9720.2006.tb02276.x>

Appendix

Original Text, Machine-Translated Texts and Task Instructions

Original text from Reisdit.no

Nice

Nice er en av de mest besøkte byene på rivieraen syd i Provence i Frankrike, og framstår i dag som en travel og livlig by. Nice i Frankrike er ett av verdens mest populære reisemål. Årsakene til det er mange. Nice er en bade- og ferieby med 8 kilometer lange flotte rullesteinsstrender.

Nice er shoppingbyen i Sør-Frankrike, fremfor noen. Her er de store motehusene representert, og her finner du småbutikker og markeder. Nice er ingen billig by. Hit har Europas overklasse og aristokrater kommet, helt siden 1800-tallet. Den fem kilometer lange vakre Promenade des Anglais, omgitt av palmer, er Nices hovedpulsåre. Her ligger legendariske Hotel Negresco, her ligger byens casino, og her finner du en mengde andre storslagne luksushoteller, eksklusive butikker og gallerier. De aller fleste hotellene langs Promenade des Anglais har eksklusive lunsjrestauranter på stranden. Prisene er stive, men en lunsj her er verd pengene. Her kan du fråse i ferske skalldyr- og fiskeretter, akkompagnert av tørre hvite franske Chablier eller Sancerres. Eller hva med å leske seg med en avkjølt lett rødvin i varmen?

Gamlebyen i Nice er et must. Du har simpelthen ikke vært i Nice, hvis du ikke har besøkt denne livlige, pittoreske delen av Nice—med trange smug, sjarmerende fargerike murhus, bitte små boutiques, en fortauskafé her og en liten intim bar der. Dump ned på en av uteserveringene ved Place de Cathédrale, titt på livet og føl atmosfæren. For en opplevelse! 1600-talls Cathédrale de Ste-Réparate er for øvrig verd et besøk—interiøret her er et godt eksempel på barokkens overdådige utsmykning. Litt nærmere klippen finner du Palais Lascaris, en 1600-tallsbygning som i dag er et museum. Her er flotte stukkarbeider, flamske billedvever, illusjonistiske himlinger og vakre og overdådige utsmykninger. Til og med et lite apotek fra 1700-tallet finner du her. Det fargesprakende blomstermarkedet La Cours Saleya, med de intense duftene er verd å få med seg, fra klokken 06.00 alle dager unntatt mandag.

Google's Translation

Agréable

Nice est une des villes les plus visités de la Riviera dans le sud de la Provence en France, et se présente aujourd'hui comme une ville animée et vivante. Nice en France est l'une des destinations les plus populaires du monde. Les raisons en sont nombreuses. Nice est une baignade et station balnéaire avec 8 km de long belles plages de galets.

La ville commerçante de Nice, dans le sud de la France, au-dessus de tous les autres. Voici les grandes maisons de mode représentés, et vous trouverez ici de petits magasins et marchés. Nice est pas une ville pas cher. Hit la classe supérieure de l'Europe et aristocrates viennent, depuis les années 1800. Les cinq kilomètres de long belle Promenade des Anglais, bordée de palmiers, est l'artère principale de Nice. Ici se trouve le légendaire Hôtel Negresco, voici le casino de la ville, et ici vous trouverez une pléthore d'autres magnifiques hôtels de luxe, boutiques et galeries exclusives. La plupart des hôtels le long de la Promenade des Anglais a des restaurants de repas exclusifs sur la plage. Les prix sont raides, mais un déjeuner ici vaut de l'argent. Ici, vous pourrez vous régaler de coquillages et de poissons plats frais, accompagnés de vin blanc sec Chabli ou Sancerre français. Ou comment doux avec un vin rouge léger refroidi dans la chaleur?

Vieux Nice est un must. Vous avez tout simplement pas été à Nice, si vous ne l'avez pas visité cette dynamique, le cadre pittoresque de Nice—avec ses ruelles étroites, de charmantes maisons colorées de briques, de minuscules boutiques, une terrasse de café et un petit bar intime il. Dump vers le bas sur l'un des jardins de la bière à la place de Cathédrale, regarder la vie et de sentir l'atmosphère. Quelle expérience! 1600 Cathédrale de Sainte-Réparate est aussi mérite une visite—l'intérieur ici est un bon exemple d'embellissement somptueux baroque. Un peu plus près de la falaise, vous trouverez Palais Lascaris, un bâtiment de 1.600 siècle qui est aujourd'hui un musée. Voici grandes moulures, tapisseries flamandes, plafond illusionniste et parures magnifiques et somptueux. Même une petite pharmacie à partir des années 1700 peut être trouvée ici. Le marché aux fleurs coloré de La Cours Saleya, aux arômes intenses sont appréciés avec le, à partir de 06h00 tous les jours sauf le lundi.

Bing's Translation

Nice

Nice est une des villes les plus visitées de la Côte d'Azur dans le Sud en Provence en France et actuellement une ville occupée et animée. Nice en France est une des destinations les plus populaires du monde. Les raisons

pour cela sont nombreuses. Nice est une ville balnéaire et vacances avec 8 km plages de galets long magnifique.

Belle ville commerçante dans le sud de la France, plutôt que de certains. Voici les grand couturiers représentés, et vous y trouverez des marchés et des petits commerces. Nice n'est pas bon marché de la ville. Voici les classes supérieures de l'Europe et aristocrates viennent, depuis le XIXe siècle. Le 5 km long belle Promenade des Anglais, entourée de palmiers, belles d'artère principale. Ici se trouve le légendaire hôtel Negresco, ici se trouve le casino de la ville, et vous y trouverez une pléthore d'autres magnifiques palaces, boutiques de luxe et de galeries. La grande majorité des hôtels le long de la Promenade des Anglais a déjeuner exclusive restaurants sur la plage. Les prix sont rigides, mais un déjeuner ici vaut l'argent. Ici vous pouvez revel dans les plats de fruits de mer et poissons frais, accompagnés de sec blanc Français Chablier ou Sancerrois. Ou que diriez-vous de doux avec un vin rouge facile fraîche dans la chaleur?

La vieille ville de Nice est un must. Vous avez tout simplement pas été à Nice, si vous n'avez pas visité cette partie animée et pittoresque de Nice-avec des ruelles étroites, charmante maison de briques colorées, petites boutiques, terrasse d'un café ici et une petite intime bar y. décharge vers le bas, dans un des jardins extérieurs à la Place de Cathédrale, regardez à la vie et l'ambiance. Ce qu'est une expérience! 17ème siècle Cathédrale de Ste-Réparate valent un visite-l'intérieur est ici un bon exemple de la somptueuse décoration Baroque. Un peu plus près de la falaise, vous trouverez le Palais Lascaris, un bâtiment qui aujourd'hui du 17ème siècle est un musée. Voici le stuc magnifiques travaux, des tapisseries flamandes, des plafonds illusionnistiske et des décorations magnifiques et somptueuses. Même une petite pharmacies du XVIIIe siècle peuvent être trouvées ici. Le marché aux fleurs colorées La Cours Saleya, parfums intenses valent pour obtenir avec lui, depuis le temps 06:00 tous les jours sauf le lundi.

Task Instructions (translated from Norwegian by the author)

Individual Work

- 1) Read the Norwegian text carefully
- 2) Examine the two translated French versions. Read through and jot down errors that you find and phrases or expressions you perceive as strange.

Group Work

- 3) Share your observations with the rest of the group. Did you mark the same problems? What type of errors and shortcomings did you find (grammar, choice of words, word order, sentence structure, text

structure. . .)? Which errors and shortcomings do you perceive as most detrimental to the understanding of the text? Why? Which errors and shortcomings do you perceive as less serious? Why?

Which of the two translated versions is the best one? Justify your choice in as much detail as possible.