This is an Accepted Manuscript of an article published by Taylor & Francis in *Language Awareness* on 28 Dec 2015 in Volume 25, Issue 1-2, 2016. Available online: http://www.tandfonline.com/doi/full/10.1080/09658416.2015.1122023

Talking about German verb particles identified in concordance lines - From spontaneous to expert-like metatalk

Johns (1997) reports that corpus-informed metatalk with a foreign language expert helps apprentice writers to make progress in independent text revision. Expecting this progress to be based on the development of expert-like ways to observe language features, I integrated Johns' so-called kibbitzing methodology into a German course based on the TV love story *Weissensee* (Hess, 2010). The course, offered to a group of French students, included online discussions focused on verb particles considered to be relevant for execution of the writing task, which was to invent a film script scene for the *Weissensee* series. The students used the specialised *Weissensee* corpus and the general corpus *deTenTen* as writing aids. Both were accessed through the *Sketch Engine* concordancer (Kilgarriff *et al*., nd). This article outlines the way in which the students talked with their teacher about language samples retrieved in concordance lines or in the first drafts of their writing tasks. Qualitative and quantitative analysis of the data - the scripts of the discussions and the transcript of the filmed interview - provides the following results: the students got more insight into the semantic complexity of verb particles, and they got a better understanding of the link between content and form.

Keywords: metatalk, concordancing; language observation; linguistic expertise; verb particles; German

1 Introduction

Tim Johns, the pioneer of the data-driven language observation method "classroom concordancing" (Johns, 1988), is known for his one-to-one kibbitzing activities[1]. These activities can be defined as corpus-informed metatalk between a student and a consultant - in fact, a teacher acting as a language expert - during foreign language text revision sessions (Johns, 1997). This method implies the use of a concordancer, a corpus exploration tool which presents all the occurrences of a search word in a listed form (cf. Table 2). It seems of interest to evaluate the extent to which this form of task-based teacher-student interaction can function as a viable support for the development of expert-like language observation. "Expert-like language observation" means using language analysis methods which can be compared to those applied by academic linguists, professional translators and - such is my claim - even by language teachers.

The usefulness of direct corpus consultation for foreign language learning has now been claimed for several decades (Johns, 1988; Aston, 2001; Braun, 2007). Learners who make corpus queries would act as "researchers" (Stevens, 1995; McEnery, Wilson & Baker, 1997) or even as "linguists" (Cobb, 2006).

Metatalk activities can take the form of what is called "metalinguistic discussions" (Gombert, 1996), "languaging about language" (Swain, 2006, p. 96) or "follow-up discussion[s] of students' use of grammar during the task" (Keck & Kim, 2014, p. 29, referring to Fotos, 2002). Their positive effects on foreign language task production are mentioned by Lyster (2007, p. 86), Suzuki and Itagaki (2007), Schoormann & Schlak (2011) and by Myhill, Jones, Lines and Watson (2012). Despite convincing results in corpus-informed language learning and in the domain of task-

[1] "One-to-one" means "one teacher speaking with one student".

based metatalk, studies which combine both of these are seldom the central issue of research papers.

The project studied in this article focuses on corpus exploitation, metalinguistic learning activities, written production and discussions about language features. Several studies explicitly mention a combination of three of these elements (Bowker, 1999; Gaskell & Cobb, 2004, Hegelheimer, 2006; Penner & Schulze, 2010; Yoon, 2008). Johns' presentation of the one-to-one kibbitzing methodology (Johns, 1997) combines all four aspects.

The *Weissensee* project includes six online kibbitzer sessions embedded in a hybrid German course. According to my working assumption, within a task-based learning setting, concordance-based metatalk can support the development of the students' capacity to elaborate adequate linguistic descriptions of language features they must master when executing the task of the project.

2 German verb particles

When you look at the linguistic features of German film scripts, you may be struck by the high frequency of verb particles. This observation is supported by the *Sketch Engine* query displayed in Table 1, checking the proportion of verb particles in the *Weissensee* corpus (Schaeffer-Lacroix, 2012), in the specialised *Geocaching* corpus[2] (Schaeffer-Lacroix, 2014) and in the general *deTenTen* corpus (Kilgarriff, Rychly & Pomikalek, 2011): the first one contains 4.98 times more verb particles than the second one and 4.25 times more than the third one.

[2] This corpus contains logbooks published by people practising an outdoor activity called "geocaching".

Table 1. *Sketch Engine*, CQL query, tag "PART.Verb".

Name of the corpus	Absolute frequency	Relative frequency (number of hits per million)
Weissensee	1,742	21,096.5
Geocaching	1,802	4,231.0
deTenTen	14,097,982	4,956.0

A certain number of prepositions have the same graphical representation as a given verb particle. In Table 2, the first concordance line contains the preposition *auf* [on], the second one the verb particle *auf*, and the third one illustrates both uses of *auf*.

Table 2. Sketch Engine, query *auf* [≈ on].

Martin wirft einen Blick	**auf**	die Uhr.	[Martin glimpses the clock.]
Martin steht	**auf**		[Martin gets up.]
Pass	**auf**	dich auf, Lisa…	[Take care of you, Lisa…]

Not only do verb particles have an ambiguous part-of-speech status, but also they tend to be highly polysemous (Krause, 2004; Kravchenko-Biberson, 2012). This explains why this article concentrates on both of these language features (verb particles and prepositions), even though verb particles are its main concern.

3 Metatalk

3.1 Definition

Metatalk can be defined as a verbal interaction between at least two people talking about selected language features. It implies (mostly written) text manipulation, cognitive activity and verbalization of the findings (Gombert, 1996). The possible cognitive effects of such metalinguistic discussions and their link with language proficiency have been discussed by Bialystok (1991, p. 116), Gombert (1996), Lyster (2007, p. 86), Suzuki and Itagaki (2007), Swain (2006) and by Myhill *et al.* (2012). The discussions can be stimulated by the learners' questions or by selected language features

considered by the teacher as significant for the execution of learning tasks.

3.2. Content - Spontaneous versus scientific metatalk

The content of metatalk can be structured with the help of the Vygotskian distinction between spontaneous and scientific concepts, mentioned in Brooks, Swain, Lapkin and Knouzi (2010). Vygotsky (2012, p. 157) evokes "the hypothesis of two different paths in the development of two different forms of reasoning". He provides details of these two forms of concepts, and he stresses the crucial role of guidance for concept formation:

> In the case of scientific thinking, the primary role is played by *initial verbal definition*, which being applied systematically, gradually comes down to concrete phenomena. The development of spontaneous concepts knows no systematicity and goes from the phenomena upward to generalization. The scientific concepts evolve under the conditions of systematic cooperation between the child and the teacher. Development and maturation of the child's higher mental functions are products of this cooperation (Vygotsky, 2012, p. 157).

A similar distinction between spontaneous and scientific reasoning can be found in Culioli and Normand (2005, p. 277): they define a *glose* as a sort of spontaneous comment about language features made by speakers who are not language experts, whereas metalinguistic comments are made by academically trained linguists aiming to formulate stable references, satisfying scientific standards of validity. The following variables support the identification of these two forms of language observation in metatalk data: firstly, the scope of the references formulated by non-experts are of more individual value, whereas those formulated by experts tend to be of more general value (cf. Vygotsky, 2012, p. 175: "the higher form of generalization known as concepts"). Secondly, language experts are able to go beyond formal description of local elements,

and they interpret the observed element by taking contextual information into account (Vygotsky, 2012, p. 260).

3.3 Method - Kibbitzing

Kibbitzing is a form of corpus-informed metatalk. This activity was first mentioned in gaming contexts. Johns (2000b) reports that during bridge tournaments, "**Kibbitzing** (looking over the shoulders of the experts)" was perceived as "instructive as well as enjoyable". According to this statement, novices benefit from kibbitzing activities at cognitive and affective levels. Within language learning contexts, this socially grounded activity consists of looking collectively - novices together with an expert - at written productions and talking about them in order to perform a writing task. Its aim is relevant and helpful interpretation of language features. Hence, it can be considered as a form of metatalk activity executed at various levels of linguistic expertise.

Johns developed a concordance-based kibbitzing service, called "One-to-one English language help" (Johns, 1997). This service was offered to postgraduate students who wanted to make progress in the area of English text production. They were given the opportunity to take a half-hour appointment with one of four consultants. The sessions aimed to help the students to become independent in terms of text revision. The consultant encouraged the student to verbalize what she (or he) intended to say in the first draft, and he showed her how to explore a general corpus or a specialised corpus in order to retrieve examples that corresponded to the revision needs. Then, the student chose among the identified examples and revised her text. She was given advice on thinking about the reasons why the language problems dealt with occurred, and on revising the same section of text at home. Then the student revised another text section highlighted as problematic and participated in another one-to-one meeting a week later.

It would be of great interest to know what was said during the meetings, but unfortunately, nobody made audio or video recordings. Johns (1997) reports that these kibbitzing sessions were felt to be very helpful by the participants, but he regrets that they did not have a strong impact on foreign language teaching. In order to make the findings of the sessions accessible to a broader audience, they were published on Johns' "Kibbitzer" pages (Scott, nd).

The *MICASE*-Team from Michigan University applies the term "kibbitzer page" to online explanations, based on language problems that have been identified in the *MICASE corpus* containing spoken academic English (Simpson, Briggs, Ovens & Swales, 2002). Reppen (2010, p. 33) calls the *MICASE Kibbitzers* (nd) "mini research projects or corpus-based language descriptions", and she explains that teachers who wish to use them for language teaching purposes need to "transform the Kibbitzer information into a ready-to-use classroom activity". As a matter of fact, the kibbitzer pages are created by researchers in linguistics, i.e. language experts. No traces of guided metatalk or other forms of individual text revision support are published on the *MICASE* kibbitzer page. It would be useful to know how students could benefit from such linguistic data.

Aston (2001, p. 1) recalls the fact that scientific (more specifically, linguistic) and pedagogical aims of corpus-use are often confused. This leads to the necessary distinction between metalinguistic discussions for and with linguists and discussions involving foreign language learners. Now, who exactly is the language expert, playing the corpus-based "revision game"? A linguist? A teacher, interested in linguistics? Any teacher? The author of the corpus data? If the principal aim of kibbitzing is language learning, the role of the expert should not only be assumed by a linguist and/or a teacher, but also by the learner: thanks to languaging activities, he may gradually

replace his spontaneous concepts by scientific ones (Vygotsky, 1986, p. xxxiv) and argue more and more like a language expert.

3.3 Triggers for metatalk

John's kibbitzer sessions aimed to meet learning needs identified in the students' written productions. The discussions were exclusively triggered by problematic text passages written by the learners. In order to push the *Weissensee* students to produce more complex texts, I completed John's reactive text revision strategy by a proactive approach (Lyster, 2007, p. 128), called "pre-emptive" by Keck & Kim (2014, p. 28). To do this, I designed focused tasks (Ellis, 2009), named "language riddles".

Table 3. Language riddle on the semantic distinction between *los, weg* and xxx.

	Los, weg oder xxx? Argumentieren Sie und kreuzen Sie die kritischen Beispiele an.		*≈ Gone, ≈ away or xxx? Find arguments and tick problematic samples.*
❑	Roman, du musst auch … .	Roman, tu dois t'en aller, toi aussi.	Roman, you too, you must go.
❑	Martin packt Falk beim Handgelenk und drückt dessen Hand … .	Martin saisit le poignet de Falk et écarte la main de celui-ci.	Martin grabs Falk by his wrist and pushes his hand away.
❑	Martin schließt den Spind … .	Martin ferme son casier.	Martin closes the locker.

The language riddles of the project invited the learners to complete worksheets, presenting pre-selected bilingual concordance lines with missing verb particles. The results obtained were the starting point for discussions about language features which were considered as relevant for execution of the main writing task (see the idea of transfer-appropriate learning, Lyster, 2007, p. 43). In such configurations, the teacher, neither too apparent nor too imposing, must strive to be available for the students during the discussions.

4 Research questions

The *Weissensee* project includes metatalk sessions for future school librarians and their German teacher. On the assumption that students can also play the role of experts during such discussions, the following research question will be examined in this paper: to what extent do the proposed metatalk activities and the teacher's interventions support the development of students' expert-like language observation?

Evidence of the learners' basic linguistic expertise is investigated by the following sub-questions: when commenting on language samples, do the students mainly concentrate on formal description? Do they use metalinguistic labels as hollow forms, without grasping the content they cover?

Evidence of high linguistic expertise is examined with the help of the following sub-questions: do the students mainly deal with meaning? Do they take into account the co-text and/or the context of the observed element? Do they provide evidence for their linguistic arguments by citing corpus examples? Are they able to explain the meaning of a sample and/or to translate it into French in an appropriate way? Do they formulate operational concepts?

5 Research methodology

5.1 The students' metatalk

The investigation of the students' metatalk was supported by a qualitative and a quantitative analysis of the transcripts of five online discussions and of the final interview. One discussion out of six, the *Skype* discussion (April 2), does not contain a sufficiently significant quantity of words produced by the learners (only 61 words in a script containing 1134 words). Therefore it is only analysed in part 7.2 which covers teacher metatalk. In order to protect the participants' identity, their first names were

changed in the discussion scripts.

The questions mentioned above were subsumed into one of six categories, ranging from basic to high language observation expertise: citing formal features, using metalinguistic labels, elaborating concepts, speaking about co-text or context, citing occurrences, and explaining or translating the observed elements (cf. 7.1). Verbal exchanges concerning management topics (how to open a file, information about the next lesson, etc.) were removed from the discussion scripts before performing the data analysis. In each of the discussions, I identified the relative frequency of the number of words of an item belonging to a given category, as exemplified in Table 4.

Table 4. Examples of word counts in the *TitanPad* discussion, May 3rd.

Categories	**Items**	**English translation**	**Number of words**	**Percentages (449 words in total)**
Meta-linguistic label	1. (là je ne vois pas bien à quoi correspond le deuxième auf, à part que c'est une particule) (…). Il est soit une particule, soit une préposition.	1. (in this case, I don't see very well what the second *auf* corresponds to, unless it is a particle) (…). It is a particle or a preposition.	26	5.79
Co-text or context	1. "Roman kippt seinen Kakao herunter" herunter oui, mais dans la phrase? 2. zu espace point: zu fonctionne avec un verbe. 3. zu permet de préciser à qui s'adresse ce signe de tête, ici au lieutenant Geifel.	1."Roman gulps down his hot chocolate" down yes, but in this sentence? 2. zu space dot: zu goes with a verb. 3. zu allows to indicate who you nod, in this case to the lieutenant Geifel.	34	7.57

5.2 The teacher's interventions

Inspired by the sociolinguistic interaction approach described in Matthey (2010), I examined the teacher's turns in the transcripts of three online discussions out of six. I selected those which, from a statistical point of view, contain enough data about teacher

interventions. I define "turn" as a verbal intervention which stops when another conversation partner begins to speak or write. The turns were classified in six categories, ranging from weak to strong teacher intervention: (re)formulating a task, agreeing, reassuring, complimenting or thanking, asking for argument, and making a linguistic statement (cf. 7.2). For each discussion, I counted the total number of turns corresponding to a given category. In order to obtain relative frequencies, this number was multiplied by 100 and divided by the number of the words contained in the discussion.

Table 5. Examples of turn counts in the *TitanPad* discussion (May 7th).

Category	**Turns**	**English translation**	**Number of turns**	**Percentages (461 words in total)**
Compliments or thanks the student	1. Intéressant – il faudra y réfléchir.	1. Interesting – it's worth reflecting.	1	0.21
Asks for argument	1. Je ne comprends pas. 2. Pourquoi pensez-vous cela ? 3. Ce contexte-là ? "Die beiden Erwachsenen nicken sich förmlich zu." 4. Oui, une précision, mais dans quel sens ?	1. I don't understand. 2. Why do you think that? 3. This particular context? "Both adults nodded formally to each other." 4. Yes, a precision, but in what sense?	4	0.86

5.3 The person marker on [one, you, we]

The observation of the frequency of the pronoun *on* [one, you, we] in all six kibbitzer discussions closes the data analysis part. The presence of this element in a linguistic statement is considered as an indicator of the fact that this statement may have a general scope (see 3.2 and 7.3 for more details).

6 The *Weissensee* project

6.1 Context and organization of the study

The *Weissensee* project was executed at the Department of Education of University Paris-Sorbonne. Three student school librarians at a pre-intermediate or an intermediate level in German were offered a 24-hour course based on the television series *Weissensee* (Hess, 2010), located in East Berlin, during the eighties, before the fall of the Berlin Wall. The course included face-to-face training in listening, reading, oral and written production. It started with a short introduction to online concordancing. The six weekly kibbitzer sessions were supported by one of the following online tools: the comment function of the learning platform *Edmodo* (Borg & O'Hara, 2008), the audio and the chat functions of *Skype* (Microsoft Corp., 2012), and *TitanPad* (Renner, 2010), a communication tool with collaborative writing and chatting features.

The students were subjected to a pre-test and a post-test inviting them to produce a short description of an animated clip containing a significant quantity of spatial events, implying the use of a certain number of verb particles.

The main task of the project consisted of writing a text which could replace or continue one of the *Weissensee* film script scenes. Film scripts contain a section called "stage directions", indicating where somebody goes or looks, etc. These text parts contain a significant proportion of verb particles. They are shown in bold in Table 6.

Table 6. Stage directions in the *Weissensee* corpus.

Stage directions	Martin stellt seine Tasche **ab**, öffnet sie und nimmt etwas **heraus**.	Martin puts his bag on the floor, opens it and takes something out of it.
Infinitive form	**ab**stellen **heraus**nehmen	put on the floor take out of

During six kibbitzer sessions, the learners and their teacher observed the meaning and the use of verb particles and/or prepositions in offline and online concordance lines

from the *Weissensee* corpus. These discussions helped the students to review the stage direction part of the first drafts of their film script scene and to produce very satisfying final texts (Schaeffer-Lacroix, 2015, to appear). A filmed interview with an external researcher in applied linguistics closed the project. I used the categories evoked in 5.1 and 5.2 as a filter when doing the qualitative analysis of the script of this interview.

6.2 Project corpora

Two online corpora were used for the *Weissensee* project: the *Weissensee corpus* (Schaeffer-Lacroix, 2012) and the *deTenTen* corpus (Kilgarriff, Rychly & Pomikalek, 2011).

6.2.1 Weissensee corpus

The *Weissensee corpus* is entirely based on the film script of the German television series *Weissensee* (Hess, 2010). This corpus contains about 82,000 tokens, and it is stored in my personal section on *Sketch Engine*[3] (Kilgarriff, Rychly & Pomikalek, nd). The data were automatically tagged with the *STTS Tagset* (Schiller, Thielen, Teufel & Stöckert, 1995/1999). I translated the first part of the series, and I created an aligned German-French *Weissensee* sub-corpus, also published on *Sketch Engine*. Each part of this bilingual sub-corpus contains about 14,000 tokens.

6.2.2 DeTenTen

The *Sketch Engine* team produced the general German corpus *deTenTen* automatically, by Web crawling. This publicly available corpus contains more than 2.8 billion tokens.

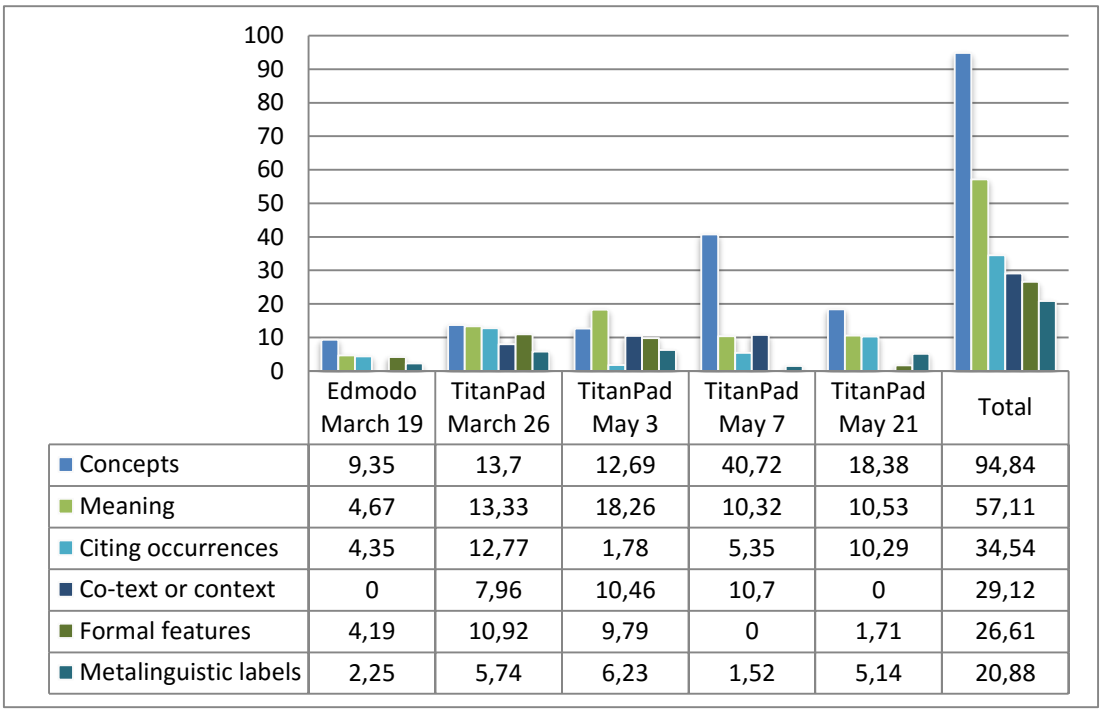
[3] The corpus management system *Sketch Engine* can be used by paying members only, but temporary access is possible through a 30-day trial option.

The students used it when they needed more data than the *Weissensee corpus* could offer.

7 Data analysis

7.1 How students verbalize their findings

In order to investigate the level of expert-like metatalk obtained by the participants in the study, five kibbitzer discussions were analysed with respect to the topics introduced in 5.1, classified from lesser to higher degrees of linguistic expertise: citing formal features, using metalinguistic labels, elaborating concepts, speaking about co-text or context, citing occurrences, and explaining or translating the observed elements. The figures in Graph 1 show the relative frequency of the total number of words concerning the topics analysed. The scripts contain between 408 and 2591 words.



	Edmodo March 19	TitanPad March 26	TitanPad May 3	TitanPad May 7	TitanPad May 21	Total
Concepts	9,35	13,7	12,69	40,72	18,38	94,84
Meaning	4,67	13,33	18,26	10,32	10,53	57,11
Citing occurrences	4,35	12,77	1,78	5,35	10,29	34,54
Co-text or context	0	7,96	10,46	10,7	0	29,12
Formal features	4,19	10,92	9,79	0	1,71	26,61
Metalinguistic labels	2,25	5,74	6,23	1,52	5,14	20,88

Graph 1. The students' topics.

According to this graph, the most frequently verbalized topic is "concepts", followed by "meaning". When describing the meaning of verb particles and/or prepositions, the

learners rarely evoke co-textual or contextual features. Metalinguistic labels and formal features are mentioned to a limited extent. The positive development of talking about content rather than simply spotting formal properties is mainly shown in the fourth discussion (May 7th) which was particularly rich in concept elaboration. Its content-oriented turns are distributed between the participants in quite a balanced manner.

7.1.1 Elaborating concepts

Table 7. Percentage of words belonging to the concept category.

Edmodo March 19	*TitanPad* March 26	*TitanPad* May 3	*TitanPad* May 7	*TitanPad* May 21
9,35 %	13,7 %	12,69 %	40,72 %	18,38 %

A closer look at the concept category reveals the fact that the learners dealt more with concepts at the end of the project than at the beginning. In parallel, they started to refer to corpus occurrences when arguing instead of communicating spontaneous impressions. To quote an example, during the very first discussion (*Edmodo* chat, March 19th), Cornélia noted that in the observed corpus examples, *an* [≈ at] would be "simpler" than *auf* [≈ on]. This statement does not help to specify the semantic differences between these two elements. In addition, she applied the concept of "immobility" (static location?) only to *an*. In fact, static location and motion events can require the preposition *an* as well as the preposition *auf*. During the chat session on *TitanPad*, May 7th, the learners evoked relevant scientific concepts such as "a precise direction", "a precise aim", "the idea of closing", "an action which causes absence", "the idea of determination", and "the will to meet the target". Only one of them – the concept of "target" – was explicitly raised by the teacher and perhaps just taken up by the students. The comparison with previous discussions clearly shows the emergence of expert-like talk about language during the project. This can be exemplified by the concept of "implication", developed when solving the language riddle (May 7th) which

invited the students to describe the meaning of *zu* [≈ to], cooccurring with *auf* [≈ on].

> Cornélia : qu'est ce qu'on entend par "implication"? Je pense au fait de participer à quelquechose par expl
> Enseignante : A votre avis, Solène et Marjorie ?
> Marjorie : Pour reprendre les exemples, c'est un peu la différence entre "écouter" et "entendre". Quand il y a implication on est plutot actif et dans l'autre cas, plutot passif.
> Enseignante : Solène, d'accord ou pas ?
> Solène: oui d'accord
> Enseignante : Oui, je veux dire avec "implication" que l'on a un lien étroit avec l'évènement ou l'élément que l'on rejoint.
> Cornélia : d'accord-donc l'idée de Marjorie du passif/actif correspond très bien
> Enseignante : Je dirais plutôt "distant" et "proche".
>
> [Cornélia: what do you mean by "implication"? I think of the idea to participate in something for instance
> Teacher: What is your opinion about that, Solène and Marjorie?
> Marjorie: If you take the examples, it is a little bit like the difference between "listen" and "hear". In the case of implication, you are rather active and in the other case, rather passive.
> Teacher: Solène, do you agree or not?
> Solène: yes I do
> Teacher: Yes, with "implication", I wanted to say that you have got a strong link with the event or the element that you are joining.
> Cornélia: okay – so, Marjorie's idea about passive/active matches very well
> Teacher: I would rather say "distant" and "close".]

Marjorie defines "implication"/"lack of implication" as being active or passive. The teacher questions this definition and states that implication has to do with the concept of distance: implication means "not being distant", lack of implication means "being distant".

Table 8. *Weissensee corpus, Sketch Engine*, "auf" combined with "zu" (right).

betreten zwei Kundinnen das Geschäft. Julia geht	**auf**	die beiden **zu**. JULIA Guten Tag.
Vera geht langsam	**auf**	die Schule **zu**. Ein paar Schüler grüßen sie.
macht sich Julia von Martin los, sie geht	**auf**	Falk **zu** und packt ihn am Kragen.
[two customers enter the shop. Julia approaches	**Ø**	them **Ø**. JULIA Good morning.]
[Vera approaches slowly	**Ø**	the school **Ø.** Some students greet her.
[Julia frees herself from Martin , she approaches	**Ø**	Falk **Ø** and grabs him by the collar.]

The corpus examples presented in Table 8 offer a context which strongly suggests that the initial distance has been reduced to nothing because the target, indicated by *auf*, has been met. This fact is expressed by the cooccurrence of the element *zu* [to, towards].

Table 9. *Weissensee corpus, Sketch Engine*, "zu" combined with "verb" (left).

das Publikum, an einzelnen Tischen sitzend, hört	**zu**	, aufmerksam und wohlwollend.
[the audience, sitting at isolated tables, listens	**Ø**	carefully and sympathetically.]

Some days later, during the *Skype session* on 14th May, Cornélia justifies her choice of *zu* in the language riddle example presented in Table 9 by taking up the concept of implication which seems to match with the meaning of *aufmerksam **zu**hören* [listen carefully]: "implication avec 'aufmerksam' "[implication with 'aufmerksam']. This statement shows that she is able to go beyond the spatial meaning of *zu* by identifying one of its more abstract properties.

7.1.2 Grasping the link between content and form

The co-text/context category, including syntactic and pragmatic properties of linguistic elements, is not very well represented in the discussions analysed. However, in the final interview, the following samples of verbal exchange about the link between content and form can be identified. The "//" signs means "speaking at the same time as the preceeding speaker", "+" means "pausing".

Sample 1

> Solène: (…) on a surtout travaillé sur les prépositions qui indiquent + qui indiquent des directions pour justement comprendre d'une manière plus fine le sens de la syntaxe allemande.
>
> [(…) we studied in particular the prepositions which indicate + which indicate directions in order to understand in fact, the meaning of the German syntax more accurately.]

By mentioning "the meaning of the German syntax", Solène establishes an explicit link between content and form.

Sample 2

> Marjorie : "contrairement au enfin au français, les les verbes allemands sont plus hein + évolutifs, enfin, je ne sais pas comment on dit, enfin,
> // Enquêtrice : Mhm. // Solène : Nuancés ?
> Marjorie : enfin, il y a une base et on peut presque lui rajouter…
> Enquêtrice : Mhm. // Solène : Oui, on rajoute.
> Marjorie : …toutes les particules qu'on veut, donc, c'est vrai que connaître le sens de la particule, c'est hein c'est plus utile pour après hein pouvoir construire les les verbes un peu à sa guise plutôt que d'apprendre des listes de verbes pour hein
> Enquêtrice : Mhm.
> Marjorie : pour en connaître le sens (…).
>
> [Marjorie: contrary to I mean to French, the the German verbs are more let's say + evolving, I mean, I don't know how you should say
> Interviewer: Mhm. // Solène: Nuanced?
> Marjorie: I mean, I mean, there is a base and you can nearly add to it…
> Interviewer: Mhm. // Solène: Yes, you add.
> Marjorie: …all the particles you want, so, it is true that knowing the meaning of the particle, you know, that is more useful for afterwards, you know for constructing the the verbs a little bit as you like rather than learning verb lists
> Interviewer: Mhm.
> Marjorie: for, you know, for knowing their meaning (…).

Marjorie says that "knowing the meaning of the particle (…) is more useful for afterwards, (…) for constructing the the verbs a little bit as you like rather than learning verb lists". It can be concluded from this statement that the kibbitzing methodology helped her to gain cognitive independence. Marjorie's description of the interplay between particles and verbs for the creation of meaning can be qualified as a sign of the development of an expert-like way to talk about language features: in fact, she provides a valid description of semantic variation.

7.2 How the teacher intervenes during the kibbitzer sessions

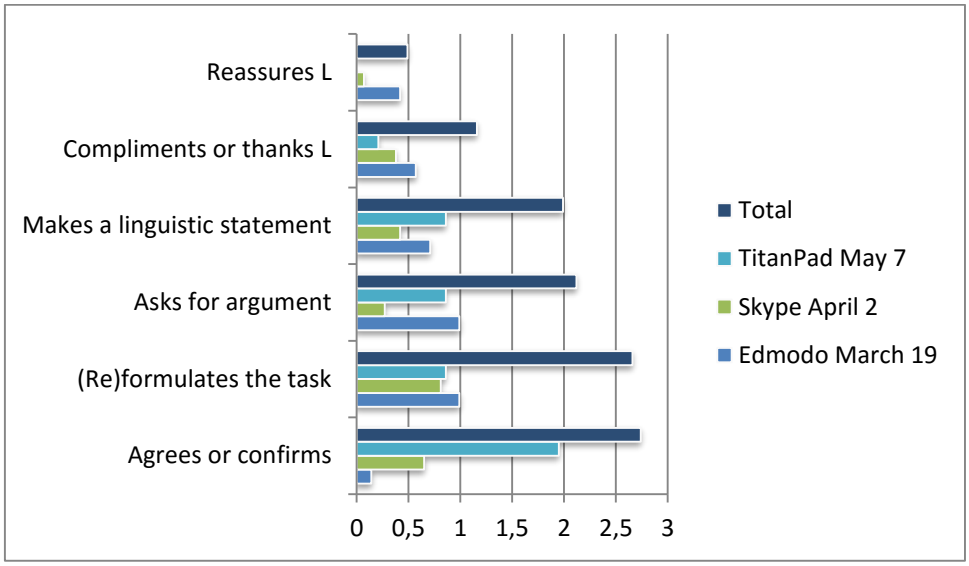
The content of the teacher's interventions, called "turns" (cf. 5.2), during distant kibbitzer sessions was classified in six categories, ranging from weak to strong intervention: she (re)formulates a task, agrees with what a student said, reassures the learners, compliments or thanks them, asks for argument, and makes a linguistic statement which completes or replaces a student's answer. The *Edmodo* discussion was mainly conducted in German, the others in French. The scripts contain between 528 and 2592 words. The turns range from 1 to 237 words.

Table 10. The teacher's turns.

	Edmodo March 19	**Skype April 2**	**TitanPad May 7**	**Total**
Agrees or confirms	0,14	0,65	1,95	2,74
(Re)formulates the task	0,99	0,81	0,86	2,66
Asks for argument	0,99	0,27	0,86	2,12
Makes a linguistic statement	0,71	0,42	0,86	1,99
Compliments or thanks the learner(s)	0,57	0,38	0,21	1,16
Reassures the learner(s)	0,42	0,07	0	0,49

The figures in Table 10 correspond to percentages indicating the relative frequency of turns, associated with categories, in the three discussions (number of turns

per number of words in each discussion).



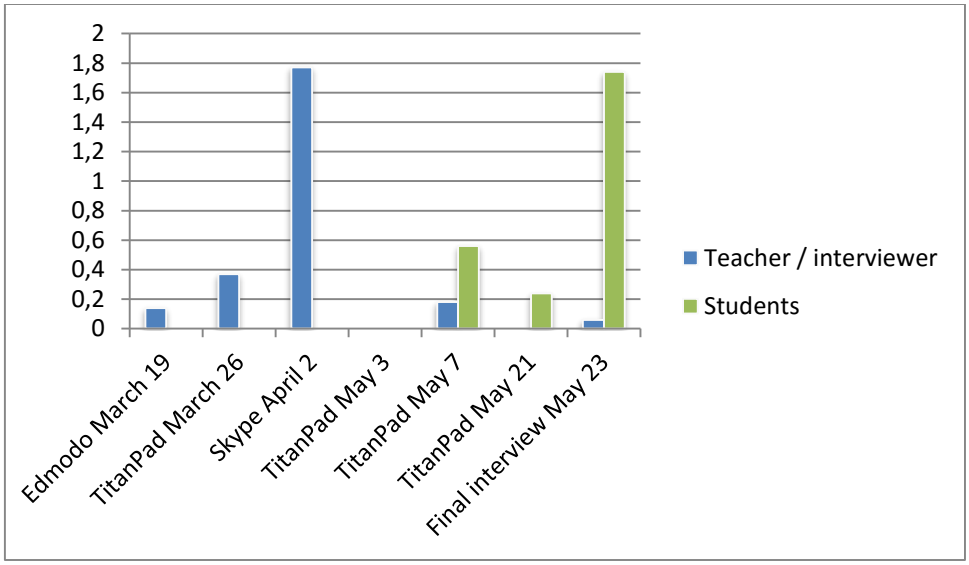
Graph 2. The relative frequency of the teacher's turns.

In Graph 2, the highest scores show up in the categories "Agrees or confirms a student's statement" and "(Re)formulates the task". The number of turns corresponding to task (re)formulation remains almost the same for the three discussions. The *Skype* discussion contains only a few invitations to present arguments; as a matter of fact, during this discussion, the teacher is mainly arguing herself. Emotion-oriented interventions such as reassuring or complimenting are decreasing, whereas agreeing or confirming the students' statements are increasing, particularly during the *TitanPad* session. This can be interpreted as a significant positive development: the students need less encouragement because their statements about language features are becoming more and more convincing.

7.3 Person markers in the discussion scripts

Elias (1968, reprinted in 1970, p. 134) uses person markers as a means to analyse the position of a speaker in relation to one or more participants in a verbal interaction. In order to observe changes concerning the scope of the students' statements during the discussions, I examined first- and second-person markers and two indefinite pronouns

(in French: *je, me, moi, vous, nous, on*; in German: *ich, dich, mich, mir, Sie, Ihnen, wir, jemand*) in the six discussions. The most relevant results concern the collective pronoun *on* [you, one, we].



Graph 3. The use of *on* [one, you, we].

Graph 3 shows the relative frequency of *on* [one, you, we] in all six discussions and in the final interview. This element dominates particularly in the teacher's interventions during the oral *Skype* discussion which has, more than the other discussions, an explicit teaching focus. In this discussion, 34 out of 46 occurrences of *on* represent an impersonal voice. The following samples shows what is meant by this term.

> Marjorie: quand on ferme une boite, on ne voit plus son contenu [when you close a box, you don't see its content any more].
>
> Cornélia: on retrouve l'idée de determination [you find again the idea of determination].

The data contain ten samples in which the speaker makes an impersonal use of *on* evoke a normative aspect, for instance, "on dit…" [you can say], "on ne dit pas" [you can't say…]. During the last but one discussion, the learners use *on* four times, representing

an impersonal voice three times. In the final interview, the students use the pronoun *on* [one, you, we] 59 times and the pronoun *je* [I] only 20 times. I conclude from these observations that finally, the learners are completing their personal voice by a collective voice. This supports the idea that at the end of the project, they argue in a more scientific way.

8 Conclusion

8.1 Results

This paper investigated the following research question: to what extent do the proposed metatalk activities and the teacher's interventions support the development of students' expert-like language observation? These are the results obtained: the execution of corpus-informed languaging activities helped the students to talk in a more scientific way about language features: they learned to go beyond formal description and to formulate abstract concepts, like "the will to meet a target". The participants developed a more fine-grained understanding of the link between content and form (Firth, 1957, p. 11). The analysis of the use of the pronoun *on* [one, you, we] illustrates the various teacher interventions during the kibbitzer sessions: after providing extreme forms of guidance, she leaves enough space for the learners, without abandoning them. The learners are thus able to formulate relevant language descriptions in a more independent way.

In the final interview, Cornélia describes the students' findings as a contribution to the research activity linked to the project.

> (…) un groupe avec un bon niveau homogène (…) permet d'entrer plus dans les subtilités. Du coup, Madame L., elle aurait peut-être pu aller encore plus loin dans toutes les particules (…). Nous (…), on est restées pas mal autour des mêmes parce qu'il y avait tellement de + variations.

> [(…) a group with a good homogeneous level would make it possible to go further into subtle details. So, Mrs L., she could have progressed even more with all the particles (…). We (…), we stayed a lot with the same ones because there is so much + variation.]

This sample suggests that thanks to the *Weissensee* project, Cornélia got more insight into the semantic complexity of verb particles: she quotes the linguistic principle of variation. In her statement, she stresses the scientific value of the students' findings: they would help the teaching language expert to "progress (…) with all the particles". Nevertheless, questions arise about the usefulness of expert-like talk about language not only for teachers and for linguists, but also for students. This could be the case in the domain of text revision. The following *Skype* discussion sample suggests that the students developed a certain form of critical distance by revising the texts of their peers during kibbitzer sessions on the *Weissensee* project: Cornélia understands that she is expected to do what is "normally" done by the teacher.

> nous devons dire avec quoi notre amie peut corriger (…) avec quel outil (…) on fait ce que **vous** faites d'habitude
>
> [we are supposed to say with what our friend can correct [her text] (…) with which tool (…) we do what normally **you** do]

As shown in Schaeffer-Lacroix (2015, to appear), this form of interaction can lead to expert-like revision and to good text quality.

8.2 Further research

I chose to analyse the online discussions of the *Weissensee* project with the help of categories inspired by previous research about expert-like language observation skills, conducted by cognitively oriented linguists or by social constructivists. As a next step, it

could be useful to analyse the "joint 'negotiation' of meaning by consultant and student" (Johns, 2000a) during metalinguistic discussions with the help of categories used in interaction research, such as those described in the third chapter of Mondada (2005) about collaborative elaboration of knowledge.

As suggested by the subtitle of the present article, "From spontaneous to expert-like metatalk", in this study I observed the development of the students' scientific reasoning capacities. Since concept formation is not a one-way phenomenon (cf. 3.2 and Vygotsky, 2012, p. 157), comparable future studies should focus on the cross-fertilization of emerging spontaneous and scientific concepts.

9 Acknowledgements

I warmly thank Annette Hess and her agent for giving me permission to use the *Weissensee* film script for teaching and research purposes. I would also like to express my gratitude to the anonymous reviewers who provided inspiring comments on a previous version of this article.

10 References

Aston, G. (2001). *Learning with Corpora.* Houston: Athelstan.

Bialystok, E. (1991). Metalinguistic dimensions of bilingual language proficiency. In E. Bialystok (Ed.), *Language processing by bilingual children* (pp. 113-140). Cambridge: Cambridge University Press.

Bowker, L. (1999). Exploring the potential of corpora for raising language awareness in student translators. *Language Awareness*, 8(3-4), 160-173. DOI:10.1080/09658419908667126

Braun, S. (2007). Integrating corpus work into secondary education: From data driven learning to needs-driven corpora. *ReCALL*, 19(3). Cambridge: Cambridge University Press, 307-328.

Brooks, L., Swain, M., Lapkin, S. & Knouzi, I. (2010). Mediating between scientific and spontaneous concepts through languaging. *Language Awareness*, 19(2), 89-110. DOI:10.1080/09658410903440755

Cobb, T. (2006). Constructivism, applied linguistics, & language education. In K. Brown (Ed.), *Encyclopedia of language and linguistics*. *Foundations of linguistics*, volume 3^2 (pp. 85-88). Oxford: Elsevier.

Culioli, A. & Normand, C. (2005). *Onze rencontres sur le langage et les langues*. Paris: Ophrys.

Elias, N. (1968). Reflections on Personal Pronouns. Paper presented at a Staff Seminar at the University of Leicester. Reprinted in Elias, N. (1970). *Was ist Soziologie? Grundfragen der Soziologie*, volume 1 (pp. 132-150). München: Juventa Verlag.

Ellis, Rod (2009). Task-based language teaching: sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221-246.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in Linguistic Analysis*. Oxford: Blackwell, 1-32.

Fotos, S. (2002). Structure-based interactive tasks for the EFL grammar learner. In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching in second language classroom*s (pp. 135-54). Mahwah NJ: Lawrence Erlbaum Associates.

Gaskell, D. & Cobb, T. (2004). Can learners use concordance feedback for writing errors?. *System*, 32(3), 301-319. http://www.er.uqam.ca/nobel/r21270/cv/conc_fb.htm

Gombert, J.-É. (1996). Activités métalinguistiques et acquisition d'une langue. *Acquisition et interaction en langue étrangère*, 8, 41-55. http://aile.revues.org/1224

Hegelheimer, V. (2006). Helping ESL writers through a multimodal, corpus-based, online grammar resource. *CALICO Journal*, 24(1), 28 pages. https://calico.org/memberBrowse.php?action=article&id=678

Johns, T. (1988). Whence and whither classroom concordancing? In T. Bongaerts, P. de Haan, S. Lobbe & H. Wekker (Eds.), *Computer Applications in Language Learning* (pp. 9-27). Dordrecht: Foris.

Johns, T. (1997). *Kibbitzing one-to-ones*. Web version of notes for presentation at BALEAP meeting on Academic Writing, University of Reading, 29th November 1997.

Johns, T. (2000a). *Tim Johns EAP Page*. Republished by Scott, M. (nd). http://www.lexically.net/TimJohns/Kibbitzer/timeap3.htm#revision

Johns, T. (2000b). *Kibbitzer.* Definition provided on Johns, T. (2000a). http://www.lexically.net/TimJohns/Kibbitzer/kib.htm

Keck, C. & Kim, Y. (2014). *Pedagogical Grammar*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Krause, M. (2004). "Konkurrenz, Komplementarität und Kooperation im Bereich der Präpositionen und Verbalpartikeln oder Wie lange noch müssen Präpositionen und Verbalpartikeln in Grammatiken ein Schattendasein führen?". *Linguistik online*, 18(1), 35-69. http://www.linguistik-online.de/18_04/krause.html

Kravchenko-Biberson, O. (2012). Le modèle cognitif et la TOE : deux points de vue sur l'identité sémantique des unités polysémiques. *CORELA - RJC Cotexte, contexte, situation*, HS 11. http://corela.revues.org/2024

Lyster, R. (2007). *Learning and Teaching Languages Through Content. A Counterbalanced Approach*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

McEnery, T., Wilson, A. & Baker, P. (1997). Teaching grammar again after twenty years: corpus-based help for teaching grammar. *ReCALL*, 9(2), 8-16.

Matthey, M. (2010). Interaction : lieu, moyen ou objet d'acquisition ?. In C. Vargas, L.-J. Calvet, M. Gasquet-Cyrus, D. Véronique & R. Vion (Eds.), *Langues et sociétés : Approches sociolinguistiques et didactiques* (pp. 31-42). Paris: L'Harmattan.

Mondada, L. (2005). *Chercheurs en interaction. Comment émergent les savoirs*. Lausanne: Presses Polytechniques et Universitaires Romandes.

Myhill, D. A., Jones, S. M., Lines, H. & Watson, A. (2012). Re-thinking grammar: the impact of embedded grammar teaching on students' writing and students' metalinguistic understanding. *Research Papers in Education*, 27(2), 139-166.

Penner, N. & Schulze, M. (2010). Group Work in a Technology-Rich Environment. *Journal for Interactive Learning Research*, 21(1), 111-137. http://www.editlib.org/p/31328/

Reppen, R. (2010). *Using Corpora in the Language Classroom*. New York: Cambridge University Press.

Schaeffer-Lacroix, E. (2015, to appear). "Impact de discussions métalinguistiques sur l'apprentissage de la production écrite en allemand, langue étrangère". *LINX* (La revue des linguistes de l'Université Paris Ouest Nanterre La Défense).

Schoormann, M. & Schlak, T. (2011). Die Unterrichtskonzeption der counterbalanced instruction. *Journal of Linguistics and Language Teaching*, 2(1), 129-168. https://sites.google.com/site/linguisticsandlanguageteaching/home-1/volume-2-2011-issue-1/volume-2-2011-issue-1---artikel-schoormann-schlak

Scott, M. (nd). *Tim JOHNS' Kibbitzers.* http://lexically.net/TimJohns/index.html

Simpson, R. C., S. L. Briggs, J. Ovens, & Swales, J. M. (2002). *The Michigan Corpus of Academic Spoken English*. Ann Arbor, MI: The Regents of the University of Michigan.

Stevens, V. (1995). Concordancing with language Learners: why? when? what?. *CAELL Journal (Computer Assisted English. Language Learning)*, 6(2), 2-10.

Suzuki, W. & Itagaki, N. (2007). Learner metalinguistic reflections following output-oriented and reflective activities. *Language Awareness*, 16(2), 131-146. DOI:10.2167/la392.0

Swain, M. (2006). Languaging, agency and collaboration in advanced second language learning. In H. Byrnes (Ed.), *Advanced Language learning: The Contributions of Halliday and Vygotsky* (pp. 95-108). London: Continuum.

Vygotsky, L.S. (2012). *Thought and Language*. Expanded revised edition. Edited by A. Kozulin. Massachusetts: MIT Press.

Yoon, H. (2008). More than a linguistic reference: the influence of corpus technology on L2 academic writing. *Language Learning & Technology*, 12(2), 31-48. http://llt.msu.edu/vol12num2/yoon.pdf

11 Corpora, resources and tools

Borg, N. & O'Hara, J. (2008). *Edmodo* [Learning platform]. http://www.edmodo.com

Hess, A. (2010). *Weissensee* [Film script of a German television series]. http://www.stichwortdrehbuch.de/drehbuch

Kilgarriff, A., Rychly, P. & Pomikalek J. (nd). *Sketch Engine* [Corpus management system]. http://www.sketchengine.co.uk/

Kilgarriff, A., Rychly, P. & Pomikalek J. (2011). *deTenTen v2.0* [German corpus stored on *Sketch Engine*]. https://the.sketchengine.co.uk/login/

MICASE Kibbitzers (nd). http://*MICASE*.elicorpora.info/*MICASE*-kibbitzers

Microsoft Corp. (2012). *Skype* [Communication tool for audio, video, chatting]. http://www.skype.com/intl/en/home

Renner, M. (2010). *TitanPad* [Collaborative online writing tool]. http://titanpad.com/

Schiller, A., Thielen, C., Teufel, S. & Stöckert, C. (1995/1999). *STTS (Stuttgart-Tübingen Tagset).* http://www.ims.uni-stuttgart.de/projekte/corplex/TagSets/stts-table.html

Schaeffer-Lacroix, E. (2014). *Geocaching* corpus [Stored on *Sketch Engine*, personal section].

Schaeffer-Lacroix, E. (2012). *Weissensee corpus* [Stored on *Sketch Engine,* personal section].

List of tables

List of figures