

# Formative assessment in English

Tony Burner

As is evident from chapter 17 in the book, assessment is a central aspect of the learning process. In this chapter, Tony Burner takes a closer look at formative assessment and how it can be incorporated in daily classroom routines. After discussing the concept of formative assessment and self-assessment, he explains the importance of formative feedback. Then he explores three forms of formative assessment, namely process writing, peer assessment and portfolio assessment.

## What is formative assessment?

As mentioned in chapter 17 in the book, formative assessment should be an integral part of teaching and learning in the Norwegian school system. In fact, for the first time assessment has become part of the curriculum (LK20). Up until Year 10, formative assessment is explained in the curriculum, providing teachers with guidelines on what matters when assessing students' products. In Year 10, there is a section on summative assessment.

There are several definitions of formative assessment, but a short one could be all assessments with the aim of improving students' learning processes and/or the teacher's teaching procedures. The aim of forma-

tive assessment is thus more and better learning for both the student and the teacher. Hence, assessment becomes a tool to promote learning. The expressions *assessment for learning* and *assessment as learning* reflect the aim of formative assessment, as is also evident in the national assessment regulations (“Vurderingsforskriften”) (Ministry of Education and Research, 2009), which have a judicial status in Norway.

However, it is important to point out that an assessment situation is not formative merely because it occurs *during* a teaching session. The Norwegian term *underveisvurdering* is often used for formative assessment, as in the curriculum and in the national assessment regulations (Ministry of Education and Research, 2009, §3-11). The term itself may be misleading in Norwegian, because it is the *purpose* that defines whether an assessment is formative or not, rather than *when* it occurs. The question is whether the assessment has the purpose to promote learning or not. A summative assessment, occurring at the end of a teaching session, may include formative purposes. A mock exam (“tentamen”) may be used formatively if the students are given the chance to reflect on the feedback provided by the teacher, revise what they have written and learn something from their mistakes. In fact, a competence aim in LK20 specifically concerns the ability to edit one’s texts (written and oral texts) according to feedback and knowledge about the English language.

A formative assessment has always an element of summative assessment, in that students’ products are assessed according to certain criteria or standards in order to be able to provide formative feedback that can help them improve their products (see Bennett, 2011). Thus, formative assessment presupposes an interaction among the students and between the students and the teacher (see Black & Wiliam, 1998, p. 16). On the one hand, the students are challenged to take more responsibility for their own learning and learning processes; on the other hand, the teacher is challenged to provide the best learning conditions for the students, based on feedback from them, so they *can* be more active and take more responsibility for their learning.

In the following, self-assessment and its formative potentials for English language teaching (ELT) will be discussed. Examples will be given of how English teachers can provide their students with opportunities to

assess their own products and learning processes. Furthermore, the topic of assessing student products will be discussed, and what it means to provide formative feedback will be explained. Process writing will be used as an example. It is not always the teacher that should provide feedback on student products. Thus, peer assessment will also be mentioned in this chapter. Finally, the use of portfolio assessment in ELT will be explained. A portfolio may contain self-assessment and peer assessment, and provides opportunities for students to be involved in their own learning processes. It is an assessment tool that can be used to differentiate teaching and learning in ELT classes.

## Self-assessment

It is a myth that children or young adults are immature when it comes to self-assessment. Large-scale review studies show that students at all ages are capable of conducting self-assessment (see Black & Wiliam, 1998; Hattie, 2009). In fact, they have a fairly precise understanding of their school achievements. When a teacher claims that students are too young or immature to self-assess, then those students are being underestimated and deprived of an important opportunity for learning. According to the national assessment regulations, students have the right to actively assess their own competence and development in all subjects (Ministry of Education and Research, 2009, § 3-12).

Why is self-assessment important? Self-assessment can make you aware of your strengths and weaknesses by reflecting on how you learn. You can become more responsible for your learning process when you have the chance to reflect on it. The quality of reflections will evidently vary from student to student, thus the teacher needs to continuously stimulate and challenge students' reflections.

Stephen Krashen's *monitor hypothesis* claims that second language learners use a language monitor to constantly edit their utterances (see Krashen, 1982). This is something many of us have experienced when we have learned new languages as adults. We learn explicit grammar rules, but when we want to speak in a natural setting, we sometimes experience that the grammar rules make us think longer and are sometimes in the

way of spontaneous speech. Krashen's hypothesis illustrates the distinction between language acquisition and language learning, where the monitor (language learning) may sometimes interfere with spontaneous language acquisition. This could be compared to the monitoring that takes place when students reflect on their learning. During self-assessment, students monitor their learning output, learning strategies and learning processes. However, research evidence points out that self-assessment is useful. In order for the language monitor not to interfere with spontaneous language acquisition, self-assessment should be conducted as a post-classroom activity. If students are asked to conduct self-assessment, which means they reflect on their learning while learning English through an activity, the language monitor may interfere negatively. Teachers may experience that students are afraid of making mistakes or are too stressed to trial-and-error with the English language. The teacher has to balance between students acquiring English spontaneously and students reflecting on their language learning.

As mentioned in chapter 2 in the book, the *European language portfolio* (ELP) can be used as a useful tool to document students' language learning and reflections (Little, 1999). The ELP consists of three components: a language passport, a language biography, and a dossier. The students make their own profile in the language passport. Their proficiency in different languages is an important part of the language passport, including a self-assessment grid describing the students' proficiency in listening, reading, spoken interaction, spoken production and writing activities. The language biography is used to plan, reflect on and assess one's own language learning and process, something that will aid in developing greater language awareness. The biography includes information about what the students can do with different languages both in and outside of school. Finally, the dossier is where the students collect evidence of their learning and development over time. However, each component of the ELP has to be adapted to the context where it is used. Teachers can use any part of the ELP. If they would like the students to self-assess their listening skills, it is possible to print the self-assessment sheet and copy it for the students. There are rubrics for each of the levels in the *Common European framework of reference for languages* (CEFR) (Council of Europe, 2001): A1 and A2 (basic user), B1

and B2 (independent user), and C1 and C2 (proficient user). The CEFR has had great impact on the development of the current curricula in foreign languages and English in Norway.

It is not only by using self-assessment grids that students can and should assess their own learning processes and products. When teachers provide feedback on a text, it can be a good idea to mark the most severe mistakes before the text is returned to the student, so that he or she is asked to reflect on and correct the mistakes (see Table 21.1). If grades are involved, it can be an advantage to let the student guess the grade before actually revealing it. It is often the case that students are stricter with themselves than the teacher. Grading one’s own work may enhance metacognitive skills, in addition to the fact that fewer students will be dissatisfied with their grades and thus there will be fewer complaints to handle for the teacher.

**Table 21.1** *Can do* descriptors for listening at B1 level.

Language:	Date for goal setting	Date for “I know this a little bit”	Date for “I know this pretty well”	Date for “I know this very well”	Date for documentation in the portfolio
Can follow a lecture or talk within his/her own field, provided the subject matter is familiar and the presentation straightforward and clearly structured.					
Can follow in outline straightforward short talks on familiar topics provided these are delivered in clearly articulated standard speech.					
Can understand simple technical information, such as operating instructions for everyday equipment. Can follow detailed directions.					
Can understand the information content of the majority of recorded or broadcast audio material on topics of personal interest delivered in clear standard speech.					
Can understand the main points of radio news bulletins and simpler recorded material about familiar subjects delivered relatively slowly and clearly					

In Figure 21.1 below, the teacher could have marked many more mistakes in the student text provided, but limits herself to pointing out repetitive mistakes (for example *their* instead of *the* and wrong use of verb tense) and expressions that can break down communication (for example the first sentence). When the student struggles with elementary knowledge of grammar, it may be counterproductive to pick on punctuation mistakes in addition to grammar mistakes.

The purpose of self-assessment and teacher assessment of student texts is not to provide students with the correct answers, but to let them go back to their texts, reflect on them and try to understand what is wrong and how to proceed with corrections. Students will have to do this on their own first, before they are guided by a more capable peer, in accordance with Vygotsky's theory on the Zone of Proximal Development (Vygotsky, 1978, p. 86).

When Mr. Clumpy below shows students their shouts and screams he's always to the children. He never speaks in a low voice or talking quietly. He is also not something good with their students and he does not listen to what they say, and he cannot be bothered to help them if they need help with a task. They other teachers have many times spoken to the principal about Mr. Clumpy, because they will give him fired. But the principal does not dare to kick him, because the principal is afraid that he will be very angry and do something dangerous. So that's why they have not kicked him out. He always said to the students that they didn't do a good job on a task. To the teachers he said a lot of ugly things, that they are not working well and that they are very ugly. When he walked in the hallways and was going to the classrooms, then he trampled and was always very heavily. So when he came to the classrooms, you could feel the floor shaking and the lookers were shaking.

**Figure 21.1** A teacher's marking of the most severe mistakes on a student text in 8th grade.

## Formative feedback

One of the most important things English teachers do, is providing formative feedback to students. Teachers want them to learn something from the feedback, and they should be given the opportunity and guidance needed to change their product according to the feedback. The way teachers provide feedback and the way the feedback is formulated can have positive, zero or negative effect on their students' learning, motivation and self-esteem. Thus, the quality of the feedback and how it is provided to the student can make a difference (see for example Hattie & Timperley, 2007). Take a look at the following scene from *Alice in Wonderland*:

“Would you tell me, please, which way I ought to go from here?”

“That depends a good deal on where you want to get to,” said the Cat.

“I don't much care where,” said Alice.

“Then it doesn't matter which way you go,” said the Cat.

This short dialogue could serve as a metaphor for formative assessment. Students have to know the next step in their learning processes. If the next step, how to improve, is unclear to them, feedback is of little help.

The task for the student text in Figure 21.1 was *Describe a fantasy teacher*. The class had prepared for the task by reading *Matilda* by Roald Dahl. In her feedback to the student who had written the text in Figure 21.1, the teacher wrote *You have used several adjectives to describe Mr. Clumpy (angry, ugly), which is good. Check if you can use even more adjectives. How does he dress?* The teacher's feedback contains information about what the student has done, in addition to pointing forward to what the student has to do next. Furthermore, the teacher focuses on just a few areas, both when it comes to language and content. The feedback is concrete, understandable and achievable.

Formative feedback is essential to formative assessment, since the whole point of feedback is learning. Teachers want the students to learn more from the feedback. Language teachers often ask themselves *How much should I correct?* They spend hours on correcting, but experience that students often make the same mistakes again. Taking into account that teach-

ers experience the students making the same types of mistakes again, it is questionable to what extent all the teacher correcting is effective (see Lee, 2009). English teachers should be supervisors for students, not editors. Teachers should ask themselves how they can guide their students to learn more and in better ways. Based on research, we know quite a lot about what types of feedback and which modes of delivering the feedback have the most positive effect on students' learning (see for example Bitchener & Knoch, 2009; Black et al., 2003; Burner, 2016; Ferris, 2003; Hattie & Timperley, 2007; Saliu-Abdulahi et al., 2017; Lee, 2011; Shute, 2008; Vattøy & Gamlem, 2020). The following advice can be given to ELT teachers:

- Think of your comments on student texts (both written and oral) as supervision or guidance.
- Make your expectations clear for the students. Your feedback should reflect your expectations.
- Provide feedback that is clear, concrete, understandable and achievable. Focus on the areas you have said you are going to focus on. “Kill your darlings” when it comes to all the other areas you have *not* said you are going to focus on.
- Point out words, expressions and sentences that may break down communication, and mistakes that reoccur, that is errors.
- Write comments both in the margin and at the end of a text, telling the student what he or she has succeeded with in the text, what needs to be improved, and, not least, how he or she can go about improving the text.
- Be sure it is clear for your students how the process of assessment works in your classes. How do you assess, why, how and when will the feedback be provided (ask them whether they prefer oral and/or written feedback), and how will they have the chance to follow up the feedback?
- Set aside time in class for responding to questions regarding the feedback and following up the feedback (for example time for revision of texts). Make sure your students understand the feedback. This is valuable investment of your time!

## Process writing

The process of writing and the final product, and the relations between them, are significant in formative assessment. Process writing is one way of highlighting these relations. In process writing, students draft, receive feedback, rewrite, revise and edit, before publishing their final products. It is not unusual that language teachers “correct” too much on local textual levels concerning orthography, morphology and syntax, and less on global textual levels like text grammar, structure and content (see Burner, 2016). Process writing requires adapted and differentiated feedback in various phases of writing. It goes without saying that the teacher should not be picky on local errors in the first draft, since this would risk killing the student’s motivation to revise or to write more.

The idea behind process writing stems from the USA and reached Norway in the 1980s. This was a period when teachers used process writing a lot with their students. The current process writing practices and perceptions in Norwegian schools have been described by Frøydis Hertzberg and Olga Dysthe (2012). Some teachers say “this was something we did in the past”, and several claim that “process writing is time-consuming”.

However, there is a high learning potential in process writing. Dysthe, one of the pioneers behind process writing in Norway, refers to international studies indicating that students at all ages revise their texts under conditions that allow them to revise. Undoubtedly, students need to practice revising their texts. Trude Kvithyld and Trygve Aasen (2012) use the term *revision competence* and underline that teachers need to instruct, supervise and model for their students how they should revise their texts. Life outside school is not any different. In fact, life outside school tends to be more authentic in that whenever you submit a text, whether it is to a publisher, newspaper editorial, poster for a campaign, post on social media, you always have the chance – and you are often required – to revise and improve it. Dysthe (1999) claims that “there is little reason to believe that feedback on finished products has any value when it comes to learning” (p. 215, author’s translation). The same is confirmed by Kvithyld and Aasen (2012): “To correct a text which is not going to be revised is ineffective response to writing” (p. 28, author’s translation). Moreover, acknowledg-

ing that a text is not finished after a first phase of writing, implies that it is unreasonable to give it a grade on a scale of 1 to 6.

There are few areas that are as conclusive as research on the impact of grades on children and young adults. Grades stimulate extrinsic motivation and competition between students rather than with oneself, particularly when it comes to low performing students (Black & Wiliam, 1998; Butler & Nisan, 1986). On the other hand, intrinsic motivation and competing with oneself have formative potential. As one lower secondary English teacher, who systematically tried out assessment without grades, said: “I have stopped giving grades together with feedback. Students work more productively if they do not receive the grades”. Research supports her claim, indicating that when students are handed back their product with a grade and feedback, they tend to overlook the feedback. The only grades the teacher *has to* give his or her students in Norway, according to the national assessment regulations, are the mid-year and final grades starting in Year 8, which is to say two times per school year. More and more schools are downplaying grades and working with improving the quality of feedback. In fact, there are some schools, particularly at the lower secondary level, that are so-called “grade free”, merely providing the two grades that have to be provided according to the regulations.

However, there is always the issue of the time aspect when it comes to feedback on student texts. It is time-consuming for teachers to provide feedback on texts and set aside time at school for the students to improve their texts, and then preferably assess the texts by giving the students a sense of the level they are at (for example “low/medium/high” or a grade). However, three points need to be mentioned. First, the teacher should express a cyclic understanding of writing through process writing, which is to say that editing and revising are a result of developing one’s ideas throughout the writing process (Hoel, 2007). Secondly, the teacher could let the students work formatively with *fewer* texts rather than not doing this with *many* texts. In that way, process writing needs not be more time-consuming than other writing activities. Instead of six-seven longer texts during a school year, a solution could be to reduce this to three-four texts which the students can work formatively with. A student who has experienced a process writing practice says:

I feel much of what I've learned in English this year comes from writing essays, working with them. I've improved my essays as much as I could. Also after having them back with comments. Actually, I've learned a lot, especially from writing many drafts.

(Korsvold, 2000, p. 113)

Thirdly, it is not always the teacher who should provide feedback on student texts. Peers can do that as well. Research from ELT classes in Norway indicates that peer assessment can actually save time for the teacher (Burner, 2015).

## Peer assessment

There are several cognitive advantages related to peer assessment, where students assess each other's products. One advantage is that students feel less intimidated receiving feedback from peers (see for example Race, 2001). This could partly be due to the fact that they are at the same age level, but also that they have somewhat similar experiences as second or foreign language learners. Norwegian teachers teaching English have been through the same processes of language learning, but in contrast to the students, those experiences go back several years in time, whereas peers often have similar questions and can help each other. It is a good idea to put students in heterogeneous groups, so that they can provide varied feedback to each other. The high performing students, too, may learn something useful by listening to other students providing and receiving feedback, and by analyzing other students' texts in order to provide feedback on them.

As already mentioned, another advantage with peer assessment is teachers saving time (see Topping, 2003). If students provide feedback on each other's products on two to five texts during a school year, it will give the teacher more time to provide feedback on all the other texts. If peer assessment is conducted regularly, systematically and professionally, teachers will experience that students often provide high-quality feedback (see Burner, 2015).

Dysthe (1999) also mentions a better learning environment, something which is essential for language learning, as one of the positive effects of a

collaborative activity like peer assessment. In a safe learning environment, students dare experiment more with the target language. Moreover, peer assessment can stimulate metacognitive learning strategies, as students learn to articulate their thinking related to learning and assessment. Peer assessment could in fact demystify the writing process for those students who avoid writing in other languages than their mother tongue, and for those students who believe writing is an ability you either have or do not have.

Learning to provide and receive feedback on language, content and structure is useful. Students reflect on their language learning strategies when conducting peer assessment. This is a transversal skill, which means that they can make use of it when conducting self-assessment at school and outside of school. In other words, there is a mutual relationship between peer assessment and self-assessment.

There are several ways of working with peer assessment at school. However, the most important thing is that students know why they take part in peer assessment, and that they are given clear guidelines on how to conduct it. If the guidelines are not clear enough, the teacher will experience that students uncritically praise each other's products without much reasoning. Here are some guidelines to the teacher regarding peer assessment, inspired by Dysthe (1999):

- After your students have written a draft, put them into groups of 3-4. Everybody in the group reads everybody's draft. The reading can take place as homework.
- Explain what the purpose of peer assessment is.
- Go through the guidelines for how to provide and how to receive feedback. Preferably, you should model a sequence of peer assessment for the class, together with a couple of the students.
- It is wise to have a set of written guidelines for the person providing feedback and for the person receiving feedback. Everybody comments on everybody's text.
- The person providing feedback has to point at some of the strengths in the text, and explain why these particular aspects of the text are positive. The examples have to be beyond the detailed local text level and

point at text grammar, structure and content. Furthermore, feedback has to be provided on a couple of aspects that are in need of improvement, and suggestions must be given on how to proceed in order to improve the text.

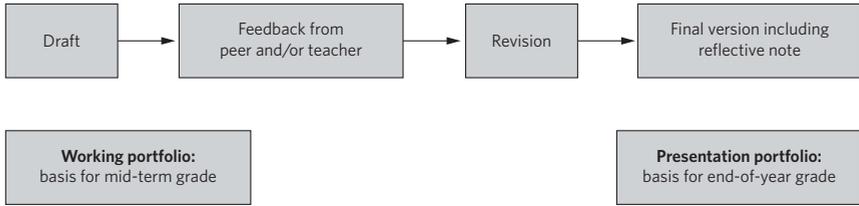
- The person receiving feedback needs to listen and take notes. It is important not to comment or try to defend oneself. He or she should be quiet until the person providing feedback is done; however, clarifying questions may be asked in order to better understand the feedback.

An MA thesis on peer assessment in English confirms the need for clear guidelines (Malesevic, 2011). The students who took part in the study were in favor of peer assessment, but continued to trust the teacher as the main provider of feedback. This also underlines the importance of practicing peer assessment. Students need to become proficient in giving feedback and to see the learning potential in peer assessment. Portfolio assessment is a system of assessment that relates the formative elements of assessment together, while at the same time preparing students for summative assessment.

## Portfolio assessment

Portfolio assessment can be regarded as an integrated approach to formative assessment where process writing is the foundation (see Klenowski, 2002). It is an assessment tool used mostly in lower and upper secondary school, and to some extent in higher education, particularly in subjects such as Norwegian and Arts and Crafts. Portfolio assessment has documented formative potentials in second and foreign language learning (Burner, 2014).

A portfolio can be defined as a systematic collection of texts that students collect, select and reflect on during an extended period. Texts do not need to be written only, but could be oral as well. The success of portfolio assessment depends on the quality of feedback, student involvement, self-assessment and peer assessment, and a focus on the process as well as the product. It is common to distinguish between a working portfolio and a presentation portfolio (see Figure 21.2).



**Figure 21.2** An illustration of how the system of portfolio assessment may look.

The former is, as the name indicates, a portfolio where students continuously document their products throughout the school year. They save drafts, receive feedback from the teacher and/or peers and revise their products. The working portfolio forms the basis for the mid-term grade, telling the students what level they are at in English. It is important to be aware that the students should also have the opportunity to revise their texts after the first term of the school year. The presentation portfolio is the final portfolio submitted for final assessment and grading. This is when the end-of-year grade is given to the students, and they will not have any more opportunities to revise or edit. Selection is an important part of the whole process, meaning that the students need to select some of the works from their working portfolio to include in the presentation portfolio, based on certain criteria from the teacher. The selected texts will be assessed according to the competence aims for the subject and/or criteria and descriptors derived from these. In order to stimulate self-assessment and reflection, students are asked to include a reflective note and attach it to the presentation portfolio. In this reflective note, students are asked to justify their selection of texts and reflect on their learning processes (see Table 21.2).

**Table 21.2** Example of criteria for the presentation portfolio.

1 non-fiction
1 fiction
1 self-chosen
1 mock exam

*Reflect on the texts you have selected for your presentation portfolio. Why have you included these texts? What are you most satisfied with? Write about the writing process. What did you learn when you worked with the texts? What did you change from first to second draft? What have you learned about writing and yourself as a writer? There are no right or wrong answers to these questions; the aim is that you reflect on and formulate your learning processes.*

No matter how English teachers conduct process writing, students need to know when the various activities take place, the deadlines for drafts and final submissions, how the teacher and the peers will assess the texts (assessment criteria), how the feedback will be provided (why not ask the students what they prefer?) and when/how they should follow up the feedback. A combination of portfolio assessment and an oral test can be useful. The teacher will have the possibility to ask questions related to the written product, in addition to the control function where you will be certain that the texts are written by the student and not by others. It could be time efficient for the teacher and systematic for both the teacher and the students to digitalize the portfolios using the school's Learning Management System. The tasks and the deadlines should be available for the students as early as possible. Feedback can also be given digitally, using MS Word's "track changes" and suitable multimedia to record oral feedback.

In ELT, it is important to reflect on one's own language learning. Elements from the *European language portfolio* can be used to form the portfolio. "I learned that I'm very good at structuring sentences, but I learned also that I could have had more content in my text", wrote one of the students in her reflective note in a research study on portfolios (Burner, 2012). The teacher had in this case provided feedback on sentence structure and given advice on what to improve when it came to the content. Another student wrote: "I learned that it is possible to improve", which summarizes well the whole idea behind a process approach to assessment (Burner, 2012).

## Concluding remarks

Research shows that feedback is more effective for students' learning than grades. Furthermore, it shows that classrooms which encourage self- and peer assessment are more effective for students' learning than those which do not, and that involving students in the assessment processes and procedures is more positive for students' motivation and learning strategies than classrooms which do not encourage and involve students in such activities (see Black et al., 2003). Research also indicates that formative assessment is useful for differentiation of teaching for every student (see Black & Wiliam, 1998). There are various ways of working with formative

assessment as an English teacher, and this chapter has given you some examples. The most important advice is not to underestimate the students and what they can achieve, but to listen to their feedback and adapt the teaching and assessing strategies accordingly and keep updated on the topic of assessment. Furthermore, it is essential to understand that formative assessment is not a certain type of activity you perform in addition to other things, but is an integrated approach within a holistic assessment culture where the value of feedback, revision, self- and peer assessment are appreciated. Assessment must never become instrumental, for example by the overuse of various assessment sheets that students feel they have to fill out but really do not understand why they have to do it or what they actually learn by it. Assessment should instead be a natural and integrated part of teaching and learning situations at school. That is when we can state that assessment promotes learning.

### Reflection questions

1. What are the most important benefits of formative assessment in English?
2. As a student, what experiences did you have with formative assessment in English?
3. What are the relations between formative assessment and transversal skills?
4. Imagine you are the student you are going to teach. How would you prefer formative assessment to be conducted in English classes?
5. What are the most important challenges of formative assessment in English, and how would you solve those challenges?

## References

- Bitchener, J., & Knoch, U. (2009). The value of a focused approach to written corrective feedback. *ELT Journal*, 63(3), 204–211.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice* 5(1), 7–71.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning. Putting it into practice*. Open University Press.

- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25.
- Burner, T. (2012). «Jeg lærer mye av mappevurdering og ser hva jeg har gjort og kan forbedre mye mer» – En studie av mappevurdering i praksis. In F. Rønning, R. Diesen, H. Hoveid, & I. Pærlussen (Eds.), *FoU i praksis 2011. Rapport fra konferanse om praksisrettet FoU i lærerutdanning* (pp. 61–70). Tapir Akademisk Forlag.
- Burner, T. (2014). The potential benefits of portfolio assessment in second and foreign language writing contexts: A review of the literature. *Studies in Educational Evaluation*, 43(4), 139–149.
- Burner, T. (2015). Processes of change when using portfolios to enhance formative assessment of writing. *Assessment Matters*, 9(2), 53–79.
- Burner, T. (2016). Formative assessment of writing in English as a foreign language. *Scandinavian Journal of Educational Research*, 60(6), 626–648.
- Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, 78(3), 210–216.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment (CEFR)*. Council of Europe. [http://www.coe.int/t/dg4/linguistic/Source/Framework\\_EN.pdf](http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf)
- Dysthe, O. (1999). *Ord på nye spor. Innføring i prosessorientert skrivepedagogikk* (2nd ed.). Det norske samlaget.
- Ferris, D. R. (2003). *Response to writing: Implications for second language students*. Routledge.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hertzberg, F., & Dysthe, O. (2012). Prosesskriving – hvor står vi i dag? In S. Matre, D. Kibsgaard Sjøhelle, & R. Solheim (Eds.), *Teorier om tekst i møte med skolens lese- og skrivepraksiser* (pp. 59–71). Universitetsforlaget.
- Hoel, T. L. (2007). Omskriving – ein kritisk faktor i skrijving og rettleiing? In S. Matre, & T. L. Hoel (Eds.), *Skrive for nåtid og framtid. Skrijving i arbeidsliv og skole* (pp. 63–76). Tapir Akademisk Forlag.
- Klenowski, V. (2002). *Developing portfolios for learning and assessment. Processes and principles*. Routledge Falmer.
- Korsvold, A.-K. (2000). *Engelsk i 10c: En fortelling fra klasserommet*. Universitetsforlaget.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon.
- Kvithyld, T., & Aasen, A. J. (2012). Å skrive, det er å omskrive – funksjonell respons for å utvikle elevenes revisjonskompetanse. *Bedre skole*, 24(1), 24–29.
- Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal*, 63(1), 13–22.

- Lee, I. (2011). Formative assessment in EFL writing: An exploratory case study. *Changing English: Studies in Culture and Education*, 18(1), 99–111.
- Little, D. (1999). *The European language portfolio and self-assessment*. Council of Europe.
- Malesevic, I. (2011). *Peer assessment as an instrument for reformed classroom assessment* (Master's thesis). Hedmark University College.
- Ministry of Education and Research. (2009). *Forskrift til opplæringslova*. <https://lovdata.no/dokument/LTI/forskrift/2009-07-01-964>
- Race, P. (2001). A Briefing on self, peer & group assessment. In *LTSN Generic Centre Assessment Series No. 9*. Learning and Teaching Support Network.
- Saliu-Abdulah, D., Hellekjær, G. O., & Hertzberg, F. (2017). Teachers' (formative) feedback practices in EFL writing classes in Norway. *Journal of Response to Writing*, 3(1), 31–55.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189.
- Topping, K. (2003). Self and peer assessment in school and university. Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment. In search of qualities and standards* (pp. 55–87). Kluwer Academic Publishers.
- Vattøy, K-D. & Gamlem, S. M. (2020). Teacher-student interactions and feedback in English as a foreign language classrooms. *Cambridge Journal of Education*, 1–19.
- Vygotsky, L. S. (1978). *Mind in society. The development of higher psychological processes*. Harvard University Press.

# Norms and variation in English language teaching

Henrik Bøhn and Thomas Hansen

According to the English subject curricula of the 1970s and 80s, English language teaching (ELT) in Norway was to be based on the native speaker norm. This meant, for example, that students were expected to learn a standard form of British or American English pronunciation. Today, however, the curriculum no longer mentions any standard for teaching and learning English. This may pose a challenge to teachers as some form of benchmark is needed for teaching and assessment. In this chapter, Henrik Bøhn and Thomas Hansen discuss the question of standards for teaching and learning English and how teachers may handle this issue in the classroom.

## The question of target language and target culture in ELT

A central idea in foreign language teaching has traditionally been that students should learn to communicate in the *target language*. Intended communication partners in this traditional approach have typically been first language speakers, and the ideal goal of instruction is for students to achieve (near-)native competence. Consequently, it has been seen as relevant to use the linguistic and cultural, or linguacultural, norms of the target culture as reference points in instruction and assessment. In other

words, the native speaker has been regarded as the norm for instruction and learning. In the case of English Language Teaching (ELT), this norm has historically been associated with British and American English, the language varieties of the two most “important” countries in the English-speaking world. The subject curriculum that came into effect in Norway in 1987, for example, stipulated that students “should [learn to use] a standard form of British or American” pronunciation (Simensen, 2014, p. 10, authors’ translation).

However, as English has now become a truly global language, it is difficult to defend the position that ELT in a country like Norway should be based solely on British and American linguacultural norms. For example, there are indications that most communication exchanges in English today are taking place without any native speaker being involved (Crystal, 2003, 2016). Curriculum developers in Norway have long since realized this, and the *English subject curriculum* in LK20 no longer makes references to any language norm for teaching and learning. Instead it stresses the role of English as a means for communication in local and global settings and the importance of *intercultural* understanding (Utdanningsdirektoratet, 2019).

Such a shift in focus poses challenges in the classroom, as some kind of standard or model is required for teaching and assessment, particularly in terms of linguistic features such as pronunciation and grammar (see for example Harlen, 2012). Learners need guidelines for knowing how to develop their language competence, and teachers need criteria for assessing it. So, if the native speaker norm is dispensed with, what kind of standard should teachers use as a model? This is a question which has been passionately discussed in the ELT community since the turn of the millennium (see Simensen, 2014).

In the present chapter, this question will be treated in some detail. Firstly, a brief historical overview of the development of English internationally will be given. Then some criticism against the native speaker norm will be presented. This will be followed by a closer look at what the current English subject curriculum says with regard to standards for learning language and culture. Finally, some practical suggestions for how to deal with this issue in the classroom will be presented.

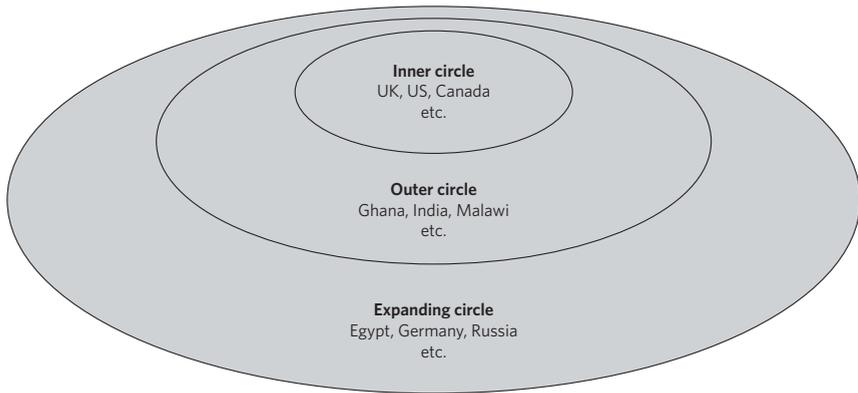
## From British English to English as a universal language

The special role of British English and American English in ELT can be explained in terms of the powerful positions of Britain and the USA internationally in the 19th and 20th centuries. The initial spread of English globally went hand in hand with the growth of the British empire from the 1700s onwards. After World War II, the extension of American economic, political and military influence around the world further increased the importance of English. Moreover, as ELT began to take hold in foreign language teaching in European countries in the late 19th and early 20th centuries, the grammars, dictionaries and phonological reference works that were used had been largely developed by British linguists and based on British English norms (see for example Howatt, 2004). In addition, the notion that ELT should be based on the linguacultural norms of the target community meant that British English became the natural choice of standard. Subsequently, American English became equally important as a reference point in English classrooms.

The extensive spread of English internationally in the 20th century meant that English came to be used in a number of new ways, and for different purposes, than was previously the case. The Indian linguist Braj Kachru (1985) has created a model representing the spread of English internationally. This model identifies three major groups of English speakers. The first group belongs to what he calls the “Inner circle” and consists of those who speak English as their *first language*. These are speakers residing in, for example, Australia, the UK and the USA. Kachru calls the countries in this circle “norm providing”, as they have typically provided speaker norms for English users outside the inner circle.

The second group has been labelled “Outer circle” and is mainly made up of citizens of the former British colonies, such as Ghana, India and Kenya, who speak English as an official or *second language*. Since many of the speakers in these countries have different first languages, English is often used as a common language of communication within their countries. Moreover, many of these speakers no longer depend on British or American or other inner circle norms, but have developed, or are in the process of developing, their own rules for how to use English. Kachru therefore calls these countries “norm developing”.

Finally, there is the group which Kachru labels “Expanding circle”, which consists of speakers in, for example, China, Norway and Russia, who speak English as a *foreign language*. In these countries English has no official status, speakers normally do not use English for internal communication, and speakers largely depend on the norms of the inner circle varieties. Consequently, Kachru has labelled the countries here “norm dependent”, since they depend on the norms of the Inner circle, in particular. Figure 22.1 gives an overview of the model.



**Figure 22.1** Kachru's concentric circles of English.

Kachru's model presents a neat overview of how English is spoken around the world, but it has not been without its critics. The British linguist David Graddol (2006), for example, has noted that the boundaries between the circles are not clear-cut, and that there are a number of speakers in the outer circle who are more proficient than many speakers in the inner circle. Similarly, he writes, there is an increasing number of language users in the expanding circle who are becoming more like second language users in the sense that their proficiency is good and that they are using English to communicate even in local contexts. Still, the model gives us a useful starting point for understanding the complexities of how English is used around the world today, and what kind of norms may be relevant for different speakers in different settings. It clearly shows that English today is much more than just British English, American English and other first language varieties. Actually, the speakers in the outer and expanding circles now

far outnumber speakers in the inner circle. The figures for the numbers of different speakers are not certain, but evidence suggests that there are more than three times as many non-native speakers of English as there are native speakers in the world today (Crystal, 2003, 2016). In other words, English has become a truly universal language, used by a great number of different speakers, in a multitude of different contexts, and for a variety of different purposes.

## Criticism against the native speaker norm

As English became ever more global in the latter half of the 20th century, an increasing number of scholars started to criticize the native speaker norm (for example Byram, 1997; Cook, 1999; Kramsch, 1998). The objections against this norm can be summarized in the following three main arguments:

- (i) It is very difficult for most foreign language learners to achieve native-speaker proficiency.
- (ii) Native-speaker proficiency is the wrong kind of competence.
- (iii) English does not “belong” to native speakers; foreign language learners have the right to use English in their own way.

The first argument is *linguistic* and rests, among other things, on research which shows that foreign language learners are rarely able to achieve a native speaker accent. Unless they start learning the foreign language at a very young age, learners will most probably always come across as “foreigners” (see Munro & Derwing, 2011). One may therefore ask why they should spend a lot of energy on trying to “sound native”.

The second argument is partly linguistic and partly cultural, and also relates to notions of identity and the purpose of learning English. As for the *linguistic* aspects of native-speaker proficiency, there is evidence that learners do not need to come across as a native speaker in order to be understood. For example, research on pronunciation and intelligibility indicates that a strong foreign accent does not automatically lead to loss of meaning (Munro & Derwing, 2011). In other words, a Norwegian learner who speaks with a Norwegian-English accent may not necessarily be dif-

difficult to understand for a foreign interlocutor. However, as will be discussed later, this depends on what kinds of “mistakes” the speaker makes. In this sense, one may say that native-speaker proficiency is the wrong kind of competence, since the overall purpose in international communication is to understand and to be understood, not to come across as native speakers (Seidlhofer, 2011).

When it comes to the *cultural* aspects of language learning, critics have argued that focusing on the cultural norms of the native-speaker culture also gives the wrong kind of competence (for example Kramsch, 1998; Byram 1997). This relates to several issues. First of all, one may ask which native speaker variety should serve as a language model in the first place. Not only are there many English first language varieties (such as Australian, British and Irish), but also within each variety native speakers are very heterogeneous, marked by their regional, ethnic, social, educational and other types of background. It is therefore problematic to choose one or two varieties as the ideal model. Secondly, as was mentioned above, English has now become a genuinely global language, where most interactions are probably taking place between non-native speakers. This relates to the third argument on the list above: English does not “belong” to native speakers, it can be – and indeed *is* – used by non-native speakers in their own ways and for their own purposes (Widdowson, 1994). Consequently, there is little reason why learners should adopt the cultural norms of, say, British English or Australian English speakers. Rather, what speakers need is to be able to interact appropriately with people of different cultural backgrounds. Or, as the *English subject curriculum* points out, an important objective of the subject English in Norway is to “develop the students’ intercultural understanding of different mindsets, ways of living and communication patterns” (Utdanningsdirektoratet, 2019, authors’ translation). No specific society or cultural area is mentioned here. For a more detailed discussion of intercultural understanding and intercultural competence, see chapter 5 in the book.

Related to the previous point is the realization that language use is also a matter of attitudes and identity. As for *attitudes* towards accents, research has shown that Norwegian students experience that Southern British English (which is more or less the same as Received Pronunciation,

or RP) signals intelligence and formality. There are also indications that British English variants are associated with prestige, or “class”. American English variants, on the other hand, are more often associated with informality and social attractiveness (see Rindal, 2010). In terms of their own accent, many Norwegian students seem to aim for either an American or a British English accent (Rindal, 2016). However, according to Ulrikke Rindal and Caroline Piercy (2013), students are not always consistent in terms of what they aim for, and how they actually speak. Students who favour an American English accent may mix British and American pronunciation, and students who prefer British English may include a number of American English features in their speech production. Still, their attitudes towards the different varieties will often influence their own choice of accent.

As for the *identity* aspect of pronunciation, it is widely accepted that accent and identity are closely related. In foreign language learning, this means that when we try to imitate first language speakers, we somehow try to “take on” their identity. However, this is undesirable to many foreign English users, as they do not want to come across as native speakers. For instance, there is evidence that a number of Norwegian students do not want to sound British or American, but aim for a Norwegian-English accent (Rindal, 2016). In this respect, they want to retain parts of their Norwegian first language identity. Therefore, the argument goes, there is no point in forcing them, since research has shown – as mentioned above – that they may be understood anyway.

## **Alternatives to the native speaker norm: English as a lingua franca**

One influential approach to global English communication which has impacted the discussion on speaker norms, is English as a lingua franca (ELF). ELF can be defined as “English as it is used as a contact language among speakers from different first languages” (Jenkins, 2015a). The above-mentioned criticisms against the native speaker norm are frequently articulated within this ELF approach. As Barbara Seidlhofer (2011), one of the “founders” of this approach, puts it:

For non-native ELF speakers, being able to use the language like native speakers and without traces of the L1 [first language] is increasingly perceived as unnecessary, unrealistic, and, at least by some, as positively undesirable.

(p. 50)

According to Seidlhofer (2011), what is important in ELF is how speakers “negotiate what is interactionally relevant, accommodate to each other, make creative use of their diverse linguistic repertoires, and cooperate in the co- or re-construction of the ‘English’ that they learnt” (p. 23). In this process English users exploit a number of different strategies and verbal (and non-verbal) resources in order to get their meaning across. The results of this are sometimes language forms which may be very different from standard (native-speaker) English. For instance, research on ELF communication has shown instances of speakers using words and constructions such as “to pronounce” (rather than *pronounce*), “examine” (rather than *examine*) and “people which I would trust” (rather than *people who I would trust*). According to Seidlhofer (2011), such words and phrases may be regarded as creative and appropriate ways to use language to communicate, rather than as language errors (pp. 102–109).

More specifically, research on language forms made by non-native speakers in ELF have typically centred on two issues: (i) lexical and grammatical features that do not normally cause misunderstandings; and (ii) pronunciation features that are important for intelligibility. As for the first point, *lexical and grammatical features*, studies have found that the following language forms are typical of non-native English users:

- using uncountable nouns as countable, for example *informations, fundings, softwares*;
- loss of 3rd person –s in present simple tense, for example *she suggest*;
- merging of who and which, for example a paper **who** will be published;
- use of all-purpose question tags, for example *isn't it* instead of *shouldn't they*;
- use of greater explicitness, for example *how long time will you stay here*;
- non-standard use of morphemes, for example *boringdom, discriminization, levelize*.

(Adapted from Jenkins, 2015b, p. 54)

As can be seen from this list, these are all language forms which would be regarded as incorrect in standard English. However, in ELF interactions, they are viewed as potentially appropriate ways of communicating.

The other feature focused on in ELF research, namely *pronunciation* and *intelligibility*, has led to the identification of a set of “core” aspects that are vital for comprehension. These aspects, referred to as the *Lingua Franca Core* (LFC), are features that learners are advised to focus on in order to make themselves understood. The main features of the LFC are:

- consonant sounds, except /θ/, /ð/ and dark /l/;
- vowel length contrasts, for example the ability to distinguish the vowels /i:/ and /ɪ/;
- not deleting sounds in the beginning and in the middle of words;
- nuclear (sentence) stress;
- the vowel /ɜ:/ (as in RP *girl*).

(Adapted from Jenkins, 2009, pp. 12-13)

As the list above shows, some pronunciation features are crucial for intelligibility, whereas others are not, such as /θ/ and /ð/. This is interesting in view of the fact that English textbooks in Norway sometimes single out /θ/ and /ð/ as important for Norwegian learners (see chapter 15 in the book). A reason for this may be that these sounds are not found in Norwegian, which means that some learners struggle to pronounce them. Nevertheless, from an ELF point of view, it is worth asking why a teacher should spend a lot of time on the difference between *thanks* and *tanks* or *then* and *den* if it is not very important for comprehensibility. Two other features which are found not to be important for understanding are intonation and word stress (for example “DEvelop” rather than “deVELOp”). Consequently, according to Jenkins (2009), they do not need to be focused on in pronunciation pedagogy either.

The advocates of English as a lingua franca point out that the language features listed above are meant specifically for ELF communication. In more traditional native-speaker focused English teaching, where the purpose is primarily to communicate with native speakers, it can be relevant to stick to the native speaker norm (Jenkins et al., 2011). What is interesting, however, is that some scholars have regarded the list of features presented above as

attempts to develop an alternative standard to the traditional native speaker norm. But this was not the intention, according to Jennifer Jenkins. The *Lingua Franca Core*, for example, was only meant to be a set of suggestions for learners to help them to speak in a more comprehensible way (Jenkins, 2015).

Two other points are interesting in this connection. The first one relates to the difficulty of completely “doing away with” the native speaker norm. Even though ELF scholars generally find the native speaker norm inappropriate, they admit that it is difficult to disregard all features of native speaker standards (see Seidlhofer, 2011). The *Lingua Franca Core*, for instance, is based on the pronunciation standards General American and Received Pronunciation. Consequently, when learners are advised to pronounce the vowel in *bird* as /ɜ:/, they are in fact being asked to use a native speaker pronunciation feature.

The other point relates to the fact that more recent ELF research has moved away from language features that may be important for understanding (such as the *Lingua Franca Core*) and has started focusing more on the diversity, hybridity, and variability that takes place when speakers of different first languages communicate in English (see Jenkins, 2015). Important key words in such communication are *appropriateness*, which means finding appropriate ways of speaking and writing, and *accommodation*, which means the ability to adapt one’s language to the interlocutor(s). The question remains, however: Which guidelines should teachers follow when they teach and assess student performance in the ELT classroom in Norway? Could the suggestions for which language features to concentrate on, made by scholars like Seidlhofer and Jenkins, replace the native speaker norm? This issue needs to be discussed in relation to the stipulations made in the *English subject curriculum*. First, however, a brief examination of what the curriculum says with regard to standards and speaker models will be made.

## **Standards, speaker models and the *English subject curriculum***

The *English subject curriculum* (LK20) emphasizes the intercultural nature of English, specifying that students should learn how to communicate with others “independent of their cultural and language background” (Utdan-

ningsdirektoratet, 2019, authors' translation). It says very little about choice of speaker model, but in the competence aims after Year 10, it stipulates that students are to “follow the rules of orthography, word inflection, syntax and text structure”. Similarly, it says that students are to “follow patterns of pronunciation”. As can be seen from the first quote, students are to follow *rules* for writing, but no specific standard is mentioned. As for the second quote, students are to follow *patterns* of pronunciation, but there is no clarification as to what kind of pattern is meant. Consequently, it may be difficult to know what kind of standard to apply when teaching and assessing English in the Norwegian context, especially when it comes to pronunciation.

## Possible ways of dealing with the question of language standards in ELT

As mentioned in the introduction, the question of models or standards in ELT has traditionally related both to the linguistic and the cultural aspects of instruction and assessment. As for the *cultural aspects* of ELT, there is a growing realization that the teaching of literature and culture related only to one or two target cultures does not fully prepare students for communication and interaction in a global world. Since there are so many cultures, both within and across nations, language users' outlook needs to be broadened. What students seem to need is to be able to deal with a multitude of different perspectives, values, identities and communication styles. One way of doing this is to develop learners' intercultural competence, as was mentioned above, in order to prepare them for appropriate communication tailored to the individual situation and interlocutor.

Concerning the *language-related aspects* of ELT, the lack of explicit linguistic standards in the Norwegian curriculum poses some challenges for teachers and learners, particularly regarding pronunciation. For example, in view of the discussion on EFL above, should teachers refrain from commenting on students' pronunciation when they say /hedetʃ/ rather than /hedeik/ for *headache*, or /tri:/ rather than /θri:/ for *three*? Or, should they accept forms such as “pronunciate”, “examinatate” or “people *which* I trust” in a written task? The answers to these questions depend on a number of factors, such as the purpose of the specific communication situation, the

genre of the task, whether the focus is on oral or written language, individual learner characteristics, and so on. When dealing with such questions, it may be a good idea to distinguish between oral and written skills.

## Suggested guidelines for dealing with oral skills

As pointed out in chapter 15 in the book, a good starting point for working with oral communication is to focus on fluency and intelligibility. *Fluency* means that teachers help students develop a smooth flow of speech without too many pauses and hesitation. This is first of all related to vocabulary, grammar, and pronunciation. As for vocabulary, no specific English norm is needed, since the goal is for students to develop as broad a vocabulary as possible. When it comes to grammar, a suitable place to start is to stick to one or two native speaker standards as *models*. In a Norwegian classroom context, where the curriculum mentions the ability to “express oneself ... using *idiomatic expressions*” (Utdanningsdirektoratet, 2019, italics added, authors’ translation), it seems safer to adhere to, say, American or Australian standard English grammar, than the suggestions mentioned by Seidlhofer, above. This does not mean that students have to follow these standards in all respects when they speak English, but that they use them as guidelines. Here it is also important to point out that learners should not be “corrected” unduly when speaking, as this may restrict their determination to speak, and thus to develop their fluency. Lastly, when it comes to pronunciation, the notion of intelligibility should be considered. As this point is comprehensively treated in chapter 15 in the book, it will only briefly be touched upon here.

A focus on *intelligibility*, then, chiefly means that teachers base their criteria for teaching and assessing pronunciation on research showing which pronunciation features that are important for understanding. As the list of core aspects (the *Lingua Franca Core*) cited above suggests, teachers can concentrate on elements such as consonant sounds, except /θ/ and /ð/, vowel length contrasts, such as the difference between *peach* and *pitch*, and sentence stress (see chapter 15 in the book for more detail). Other features, such as intonation, weak forms and assimilation, may be less important (Jenkins, 2009).

However, there is evidence that many learners actually do want to model their spoken production on native speakers (see for example Rindal, 2016). Consequently, teachers should be prepared to help these learners achieve that goal. Yet, there is little justification in the *English subject curriculum* for insisting that all students should aim for a native speaker accent. This needs to be kept in mind when developing criteria for final assessment. It is very difficult to justify the claim that a student would have to speak like an American or a Briton in order to get a 6 in the oral exam. A candidate who has excellent vocabulary, flawless grammar, very good fluency and so on may still be awarded the top grade even if his or her Norwegian accent shines through.

An additional point regarding pronunciation in ELT is that teachers always need to consider different factors, such as individual learner characteristics, the purpose of the communication or learning situation, intended interlocutors and so on. In this respect, the notions of *appropriateness* and *accommodation*, mentioned above, are key principles for suitable language behaviour. Rather than native speaker perfection, then, a focus on intelligibility, fluency and the students' ability to adapt to the situation and the interlocutor, is a better starting point for developing learners' English oral skills. These aspects can also serve as principal guidelines for assessing performance, both formatively and summatively. Students who aim for a native speaker accent should of course have the chance to do that, and teachers can provide resources in the form of authentic audio material, online pronunciation guides, learner's dictionaries and so on. Teachers should also be prepared to supervise these learners to help them improve further. Yet, the main focus should be on general communication skills and the ability for students to interact appropriately in different settings (for further reading, see chapter 15 in the book).

## **Suggested guidelines for dealing with written skills**

In terms of *writing* instruction and assessment, many of the points mentioned under oral skills also apply here. As the basic skill relating to written proficiency in the *English subject curriculum* states, a major concern

in the development of writing abilities is to enable students to “express ideas and meaning in a comprehensible and purposeful way in different types of texts” (Utdanningsdirektoratet, 2019, authors’ translation). Again, this emphasizes the students’ ability to tailor their communication to each specific situation that they encounter. However, this does not mean that accuracy, or grammatically correct language use, is irrelevant. As was touched upon above, the curriculum also specifies that students need to adhere to language rules and to express themselves in idiomatic English. All the same, the failure to mention which language norm these language skills are to be modelled on makes it difficult for teachers to know which standard to use.

From a pragmatic point of view, it may be appropriate to use Standard written British English, or Standard written American English, as a basis when teaching written communication in English (see for example Haselgård et al., 2012). These two varieties are very well described in grammars and learning materials, and they have so many features in common that they are not likely to confuse many learners. There is also a long tradition in Norway of using these varieties, which means that teachers, learners, and the general public are very well acquainted with them. Moreover, they are much used internationally in both professional and non-professional settings, which speaks in favour of their usefulness.

That said, it is important to note that the use of these standards does not mean that one cannot include references to other standards, if relevant. Also, as students are to be able to use English in various settings and for various purposes, they *should* be made familiar with non-standard varieties and how they are, or can be, used in different contexts. Flexibility and the ability to accommodate are important. Still, in a school context, it can be pedagogically sound to take Standard English usage as a starting point.

As for the assessment of written proficiency, this, of course, depends on the purpose of the writing and the genre. When working, for example, with memes, tweets or other types of informal, electronic texts, it may be very appropriate for students to use informal language, slang, abbreviations and even ungrammatical forms. In more formal texts, however, a phrase like “people which” should not be accepted, and students need to be made aware of the more correct, standard form.

## Concluding remarks

The question of norms, standards, models, and guidelines in ELT is a difficult one. In today's global community, where English is used more outside than inside the "inner circle", and where appropriate communication is typically the ultimate goal, there is little point in expecting students to speak and write like native speakers. However, some reference points are needed in language education, and the rejection of the native speaker norm in many ELT contexts has left the question open as to which standard should be used. The *English subject curriculum* reflects this uncertainty. A possible solution to this problem is to use some form of native speaker standard as a *starting point* and a *model* for teaching and assessment, but with a clear understanding that the purpose and the situation of the communication situation should always decide how strictly one needs to stick to the model. Generally, it may be a good idea to focus more on a native speaker standard in formal written communication, whereas intelligibility may be a better focal point in most forms of oral communication.

### Reflection questions

1. To what extent were you expected by your English teachers to approximate a native speaker accent when you learnt English in school?
2. What are the advantages and disadvantages of using the native speaker norm as a basis for English language teaching (ELT)?
3. Why has the intelligibility principle gained ground in ELT in later years?
4. How can you be good role models to your students in terms of pronunciation?
5. What kind of pronunciation features would you focus on in the English language classroom? Is it necessary to correct students who say /tɪŋk/ rather than /θɪŋk/, /fast/ rather than /fɜ:(r)st/, /ɪv/ rather than /li:v/ and /'devələp/ rather than /dɪ'veləp/?

## References

- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters.
- Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly*, 33(2), 185-209.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Council of Europe. [http://www.coe.int/t/dg4/linguistic/Source/Framework\\_EN.pdf](http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf)
- Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge University Press.
- Crystal, D. (2016). *Who would of thought it? The English language 1966-2066*. Paper presented at the IATEFL annual conference, 13 March, Birmingham. [http://www.davidcrystal.com/?id=-1592&taxonomyselctid=-1487&report\\_\\_pageno=9](http://www.davidcrystal.com/?id=-1592&taxonomyselctid=-1487&report__pageno=9)
- Graddol, D. (2006). *English next: Why global English may mean the end of 'English as a foreign language'*. The British Council.
- Harlen, W. (2012). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning* (2nd ed.) (pp. 87–101). Sage.
- Hasselgård, H., Johansson, S., & Lysvåg, P. (2012). *English grammar: Theory and use* (2nd ed.). Universitetsforlaget.
- Howatt, A. P. R. (2004). *A history of English language teaching* (2nd ed.). Oxford University Press.
- Jenkins, J. (2009). (Un)pleasant? (In)correct? (Un)intelligible? ELF speakers' perceptions of their accents. In A. Mauranen, & E. Ranta (Eds.), *English as a lingua franca: Studies and findings* (pp. 10-36). Cambridge Scholars Publishing.
- Jenkins, J. (2015a). *Global Englishes – A resource book for students*. Routledge.
- Jenkins, J. (2015b). Repositioning English and multilingualism in English as a lingua franca. *English in Practice*, (2)3, 49–85.
- Jenkins, J., Cogo, A., & Dewey, M. (2011). Review of developments in research into English as a Lingua Franca. *Language Teaching*, 44(3), 281–315.
- Kramsch, C. (1998). *Language and culture*. Oxford University Press.
- Kachru, B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk, & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge University Press.
- Munro, M. J., & Derwing, T. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching*, 44(3), 316–327. <https://doi.org/10.1017/S0261444811000103>
- Rindal, U. (2010). Constructing identity with L2: Pronunciation and attitudes among Norwegian learners of English. *Journal of Sociolinguistics*, 14(2), 240–261. <https://doi.org/10.1111/j.1467-9841.2010.00442.x>

- Rindal, U. (2016). "Riktig" engelsk uttale. *Bedre skole, 1*, 79–84. <https://utdanningsforskning.no/artikler/riktig-engelsk-uttale/>
- Rindal, U., & Piercy, C. (2013). Being 'neutral'? English pronunciation among Norwegian learners. *World Englishes, 32*, 211–229.
- Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford University Press.
- Simensen, A. M. (2014). Skolefaget engelsk. Fra britisk engelsk til mange slags "engelsker" – og veien videre. *Acta Didactica Norge, 8*(2), Art. 15. <https://www.journals.uio.no/index.php/adno/article/viewFile/1138/1017>
- Utdanningsdirektoratet. (2019). *Læreplan i engelsk*. <https://www.udir.no/lk20/eng01-04>
- Widdowson, H. G. (1994). The ownership of English. *TESOL Quarterly, 28*(2), 377–389.



# Corpora in English language teaching

Hilde Hasselgård

Corpora, in the sense “databases that contain texts”, are a widely used tool in language research. This tool can be a helpful resource for language teachers and learners as well, and many corpora are available online. In this chapter, Hilde Hasselgård discusses the use of corpora in the English classroom. She explains what a corpus is, and what answers a corpus can provide to questions about language use. Then she presents various possibilities for making use of corpora in language learning.

## Finding out how language is used

Linguists use corpora to investigate how language is used in natural texts and to test their hypotheses about language. Such investigations sometimes concern the frequencies and the patterns of words. What words are used more or less frequently than others and in what types of texts? What lexical and grammatical patterns do words tend to occur in? What words and grammatical patterns tend to occur together?

Investigating language through corpora is not limited to researchers: A corpus can be a useful tool for language teachers and learners alike. In a language learning situation a corpus can act as a reference tool along with dictionaries and grammars. Everyone can search for particular words and expressions and get examples of how these are used in natural contexts.

This approach to language is descriptive rather than normative: It shows people how language is used in actual discourse but does not tell them how to speak or write. Unlike a dictionary, a corpus does not explain what a word means, but it can show how the word is used in sentences. It thus enables corpus users – after having interpreted the search results – to answer questions such as:

- What preposition is used after *critical*?
- Do people say *eggs and bacon* or *bacon and eggs*?
- Is the word *bloody* still used in the sense “full of blood”?
- Can I write *and so on* in a formal text, or is *etc.* more common?

Many corpora are freely available on the Internet. Some of these are listed at the end of this chapter. Corpus linguistics is a fast-developing field, so there may be good and useful corpora that are not listed. Nevertheless, the list should give plenty of material for studying speech and writing in the major varieties of English.

## What is a corpus?

As explained above, a corpus is a digital collection of texts. A more precise definition is that a corpus is a structured database of natural texts prepared for use in linguistic research. That is, the texts are not produced in order to be included in a corpus, but in order to communicate in authentic settings. However, the selection of texts and the way in which they are organized have been planned with linguistic research in mind. This means that not all databases of texts are corpora, because they may have been created for other purposes. Examples are archives of newspapers and public documents. When linguists compile a corpus, they typically aim to make the corpus *representative*; that is, it should give a fair image of the language of a particular period, region or genre, for example. Users of a corpus need to be aware of what the corpus contains and thereby what kind of language it is meant to represent. For example, we cannot make claims about language in literature or in classroom settings based on a corpus that consists exclusively of newspaper text. Similarly, a corpus that contains only writ-

ten material cannot show us how people speak, and one that contains only conversations cannot reveal whether the singular *is* in *There's hundreds of tourists* is acceptable in writing.

The fact that a corpus has been prepared for use in linguistic research means, among other things, that the end user can see where the texts come from. In the *British National Corpus* (BNC), for example, each sentence of each text has a code that reveals its source. So in the case of the sentence below, the code (KS8 785) reveals – if we click on it in the corpus display – that the sentence comes from a concert programme.

Tonight's programme features music from all these styles. (KS8 785)

In the BNC, as in many other corpora, the preparation also includes *tagging* each word for part of speech, which means that each word gets a tag attached to it with a code for the word class it belongs to. Such tags are thankfully not visible in a normal view of the corpus, because they clutter the text spectacularly. Here is what the above sentence looks like with visible tags:

Tonight<AVo>'s<POS> programme<NN1> features<VVZ>  
music<NN1> from<PRP> all<DTo> these<DTo> styles<NN2>.

However, because the tags can be included in corpus queries, the searches can be made very precise. For example, we can search for *feature* as a verb (as opposed to a noun) or for an adjective (such as *clever*) followed by any preposition. The latter search in the BNC shows that the most frequent preposition to occur after the adjective *clever* is *of*, as in the first sentence below. Next in frequency we find *at*, *by*, *with* and *for*, as in the next sentences below. These sentences illustrate not only different expressions with *clever* and a preposition, but also different meanings the adjective.

It was very *clever of* you to find it.

He was *clever at* finding bargains.

Oh, you're too *clever by* half.

Still, I expect she's *clever with* her hands.

According to her, he was too *clever for* his own good.

## Concordance

A corpus may come with a so-called search interface. That means that it has its own search tool which is tailored to searching in the corpus and displaying the output in a way that is useful for language study. This leads us to another central concept in corpus linguistics: *concordance*. A concordance is a list of the occurrences of the word or phrase that was searched for. Each occurrence has a context, which can be a sentence or a specific number of words or characters on either side of the search word(s). Figure 23.1 shows what a concordance may look like. The search word, *hopeful*, is highlighted and appears in the middle of the line. The lines have been sorted so that they appear in the alphabetical order of the word to the right of *hopeful*. This function of the search interface makes it easier to see what patterns the word occurs in. For example, the concordance in Figure 23.1 shows that *hopeful* can describe both people and things, with the difference that a “hopeful person” is hoping for something, but a “hopeful thing”, as in *a hopeful sign*, is one that gives people hope. The concordance also gives an example of *hopeful* used as a noun. We can compare *hopeful* with the similar adjective *wishful*, which turns out to have a much less varied pattern of use in the same corpus: It occurs only in the expressions *wishful thinking* and the related *wishful thinker*.

<input type="checkbox"/>	Details	Left context	KWIC	Right context <span style="font-size: small;">↓</span>
1	<input type="checkbox"/> ⓘ Christian Unity in ...	:hurches . </s><s>	Many people seem <b>hopeful</b>	, yet it is difficult to predict whether or n
2	<input type="checkbox"/> ⓘ Land of the Silver ...	istance , while the men were stern but	<b>hopeful</b>	. </s><s> All , of course , except the Dc
3	<input type="checkbox"/> ⓘ various	tion of the seven principal Presidential	<b>hopefuls</b>	: five Democrats -- Senator Hubert H. F
4	<input type="checkbox"/> ⓘ Deadlier Than the ...	s><s> Jeb cautioned him not to be too	<b>hopeful</b>	and then , ignoring his own advice , sai
5	<input type="checkbox"/> ⓘ My Hero.	ik Adam Herberet is guilty of being too	<b>hopeful</b>	and better informed on defense financi
6	<input type="checkbox"/> ⓘ The Heartless Light.	y else to tell them : no assurances , no	<b>hopeful</b>	hints at great discoveries that day . </s
7	<input type="checkbox"/> ⓘ Values and Moder...	rtising agency ; ; </s><s> and many a	<b>hopeful</b>	incipient business executive decides it
8	<input type="checkbox"/> ⓘ various	ig March 4 . </s><s> Mr. Hodges is so	<b>hopeful</b>	over the outlook that he does <g> n't ti
9	<input type="checkbox"/> ⓘ Peace Corps. Fact...	s new peaceful program , this will be a	<b>hopeful</b>	sign to the world . </s><s> Congress si
10	<input type="checkbox"/> ⓘ A Passion in Rome.	it aside for good . </s><s> But it was a	<b>hopeful</b>	sign , he told himself . </s><s> She no

**Figure 23.1** Concordance for *hopeful* (Brown Corpus, via Sketch Engine).

## Frequency

Corpus linguists are often interested in *frequency*; that is, how often a word or expression occurs in the corpus, particularly in comparison with other words and expressions. Information on frequency is useful in choosing among alternative wordings. For example, the word *different* can occur with various prepositions: *from*, *to* and *than*. A search in the BNC for *different* + preposition shows that *different from* occurs 3243 times, *different to* 429 times, and *different than* 50 times. Thus the safest option is *different from*. It is of course possible that the prepositions signal different meanings, as we saw with *clever* above. This is also the case with *sorry for/about*, as indicated by concordance lines with those combinations in the BNC (in Table 23.1).

**Table 23.1** *Sorry for and sorry about* in the British National Corpus.

H8T	But when nobody else feels sorry for you, you tend to feel <b>sorry for</b> yourself, don't you?
KBL	No I'm not being rude to Brian! I feel <b>sorry for</b> Brian! What! I'm really hurt by that! Well that's
HWL	Swedish number plates in order to follow him. I felt a wee bit <b>sorry for</b> the driver—reindeer probably don't drive like that—but
CDM	. I'd liked her until the money lending began, and I was <b>sorry for</b> being unkind to her. Then Frankie saw me. " No use
G06	horrible but who is too polite to draw attention to the fact by seeming <b>sorry for</b> her. I didn't yet know Lili.
CDM	went out dolled up. Everybody seemed to shun her but I felt <b>sorry for</b> her and we became close friends. However, she didn't
HHC	, I know you've said it doesn't matter, but I am <b>sorry about</b> what happened up at Handley Farm. " Sorry for what,
CCE	make do with animal furs and leather where we are going. ( <b>Sorry about</b> that.) Let's pause for a while in order to breathe in
HGM	this time by Ace. " Dara, great to see you! <b>Sorry about</b> last night! Do you two girls know each other? " Kate, on
B0U	which comes from solitary confinement was already maturing inside me. " <b>Sorry about</b> the bread. It was the best I could do.
FP7	; if it comes to that, so do you. " I'm <b>sorry about</b> the phone call. I got the instruction but I didn't really take

In order to compare frequencies from two different corpora, we need to know how big the corpora are. In the case of *different from*, for instance, we find 35 examples in the Brown Corpus, which was used for Figure 23.1, while the BNC has 3243. But since the Brown Corpus contains one million words and the BNC has 100 million, *different from* is actually slightly more frequent in Brown, with 35 vs. 32.4 occurrences per million words.

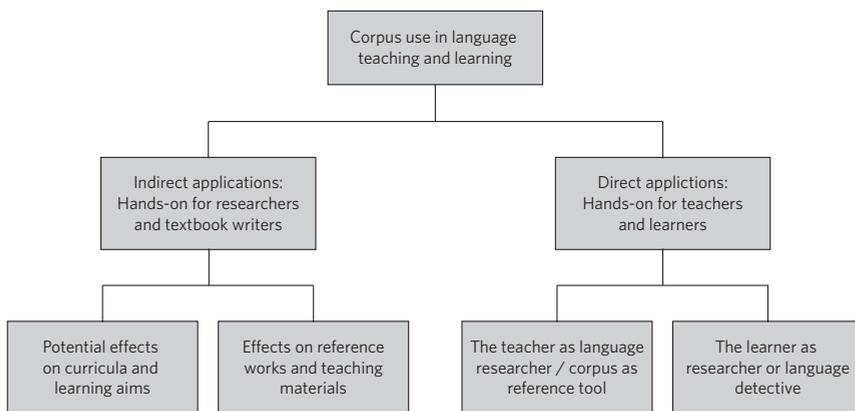
## Variation

Corpora can also reveal *variation* in language use. The variation can be between different patterns of the same (or similar) words, as in the examples with *different from/than/to* and *hopeful* vs. *wishful*. But it can also be between regional varieties of English or different text types. The corpus search interface at <https://www.english-corpora.org/> lets the user compare text types. For example, a search for the word *stuff* in the *Corpus of Contemporary American English* (COCA) shows that the word is clearly most common in speech and rather rare in academic writing. Using the GLOWBE corpus (short for *Corpus of Global Web-Based English*) we can find, for example, that the modal auxiliary *ought to* is more frequent in Nigerian English than in any of the other regional varieties included in the corpus, and least frequent in Irish English. This information may not be of much interest in itself, but could inspire further investigations of modal expressions in different varieties of English.

Another aspect of the linguistic variation in a corpus is that not all corpus examples will be universally recognized as good or correct usage. This is an inevitable consequence of the fact that the corpus consists of authentic texts. Speakers/writers simply do not always express themselves in ways that language learners should copy. Therefore it is a good idea to use frequency as a pointer to what expressions are preferred in the corpus, since we can assume that acceptable uses will outnumber erroneous ones. When in doubt about the acceptability of a corpus example, we can also look up the same expression in a dictionary.

## How can we use corpora in English language teaching?

The answer to this question depends on a number of factors, especially the proficiency level of the learners, their learning goals, and the practical issue of whether computers are available in the classroom. Corpora can be helpful to a language teacher regardless of how these factors play out, but the degree of involvement with the corpus will vary. Various uses of corpora in language teaching and learning are visualized in Figure 23.2.



**Figure 23.2** Corpus use in language teaching and learning (based on Römer 2011, p. 207).

### Indirect applications of corpora

Even without using corpora ourselves, we can benefit from teaching materials and reference tools that are based on corpus investigations. For example, most major English-language dictionaries are corpus-based. As a result, the dictionaries can give reliable information on the kinds of contexts a word or phrase occurs in; for example whether it is academic or colloquial, rare or frequent. Besides being an excellent source of authentic examples that can show the usage patterns of words and phrases, corpora can be used to define a core vocabulary. Studies have shown that the 2000 most frequent words in a ten-million-word corpus of spoken and written English account for 83% of the text (O’Keefe et al., 2007). This means that a learner who has

acquired these words will get by in most situations. In addition, it may be helpful to work with specialized corpora to identify and teach slightly less frequent words that are nevertheless useful in the kind of situations and genres that the learners are likely to encounter. One use of specialized corpora is to extract word lists to tailor the vocabulary teaching to specific learner groups and learning purposes. For example, young learners may focus on vocabulary that is frequent in corpora of children's books and everyday conversations in family settings, while those learning for a particular profession (such as cooking, carpentry, engineering, business) could learn vocabulary drawn from corpora of books, magazines, journals and transcribed lessons within their field, or, say, from recorded service encounters if the goal of the learning is to be able to serve customers in English.

Grammarians have drawn on corpora to improve the description of language structure. Here, too, the corpus is an invaluable source of authentic examples of grammar in use. But even more importantly, the corpus can provide accurate information on the occurrence of particular grammatical constructions in different contexts. An example of a corpus-based English grammar is the *Longman Student Grammar of Spoken and Written English* (Biber, Conrad & Leech, 2002). This book contains information that would have been impossible to discover and document without the corpus, for example that the passive construction with *get* is most common with the verbs (*get*) *married*, *hit*, *involved*, *left* and *stuck*. Other verbs seldom make use of their potential to occur in the passive, for example, *exclaim*, *guess*, *hate*, *joke*, *love*, *try* and *want* (pp. 171–172). The corpus that this grammar is based on contains four genres: news articles, fiction, academic prose and conversation. We can thus also learn that the passive is most common in academic English and least common in conversation (p. 178). Less obviously, the passive is more common in news than in fiction.

Corpora can furthermore be the basis of exercise materials for both vocabulary and grammar learning. It is possible to find ready-made exercises of this kind both online and in printed form, often under the heading “data-driven learning”. However, it is also relatively simple for teachers to design their own exercise materials from a corpus. For example, we can create exercises on “confusable words” (for example *teach/learn*, *take/bring*) by searching for these words, deleting the “keywords” from the concordance

lines so that the students can fill them back in, as shown in the sample exercise. It may be useful to omit concordance lines from the exercise that might confuse the learner, for example because of difficult vocabulary. If students have corpora available to them, they can create such exercises for each other.

### Sample exercise

This exercise is based on a selection of sentences retrieved from searches for *teach* and *learn* in the British National Corpus. In the exercise, these words have been removed. The same technique can be used for other word pairs or for different grammatical forms (such as *walk/walks, walks / is walking*).

Instruction for the learners: Fill in *teach* or *learn* in the sentences below.

1. And our spinners too have had to \_\_\_\_\_ how to bowl on these pitches.
2. He goes on to argue that we can \_\_\_\_\_ to cope with the anxiety associated with an anticipated event.
3. It may not be possible to get rid of the mite completely but we can \_\_\_\_\_ to live with it.
4. It will also help \_\_\_\_\_ your child computer interaction at an early age.
5. Keep an eye on how employees develop, and \_\_\_\_\_ from their mistakes.
6. She had the joint X-rayed on Wednesday night and was somewhat relieved to \_\_\_\_\_ that it was merely a bad case of bruising.
7. That will \_\_\_\_\_ your father to allow your "admirers" to visit the Black Lion.
8. We have to \_\_\_\_\_ from the past.

## Direct applications of corpora: teachers and learners as corpus users

Regardless of the proficiency level of the students, a language teacher can use corpora to check words and expressions in student texts, to create teaching materials, and to find good illustrations to explain words, phrases and grammatical constructions. A big corpus will help a non-native English teacher answer the question *Can you say this in English?* It can also be

used to check a suspicion that the style level is inappropriate. For example, learners commonly use a lot of informal and spoken-like expressions in their written English. A corpus can show us that certain expressions belong mainly to conversational style, such as *you know, a little bit, really great*. Similarly, we can search for expressions that sound odd. If they cannot be found in the corpus, or there are very few of them compared to other expressions with similar meaning, we can use this to advise learners to choose different wordings. However, it is important to remember that no corpus can contain everything. The fact that an expression is not found in the corpus is not proof that it does not exist in the language. A word or expression may, for instance, be more recent than the corpus texts, or it may belong to a genre that the corpus does not contain.

An exciting feature of corpus use is that students can use the same methods as language researchers. Even at relatively low levels of proficiency, learners can come to the corpus with questions about the language and discover something new. And as Stig Johansson (2011) observes, corpus users often make interesting discoveries in addition to what they were originally looking for. In order for learners to benefit from corpus use, however, it is important that the material is suitable for their level. Most language teachers have experienced that learners at low levels of proficiency are overwhelmed when faced with authentic texts. It is therefore vital that students are given manageable tasks. As they become more proficient, they can handle increasingly more advanced and independent tasks.

When exposing young learners to corpus material, it may be useful to give them a simple concordance focusing on a word or phrase that is central to what the class is working on. The sentences below came out of a search for *cat* in an interface to the BNC where it is possible to specify the target readership of the texts, in this case children and teenagers (unfortunately this interface is not freely available, but see the section on “Do-it-yourself corpora”). Such sentences can be the starting point for work with words and expressions for describing a cat and its activities.

So the striped **cat** made her way to the garden of the tower. (FUB 611)  
 Slowly the **Cat's** eyes, then its ears, and then the rest of its head  
 appeared. (FNS 448)

The **cat** was black and white: half its face was black and half was white; half its body was black and half was white. (FSL 251)

Tom **Cat** jumped down. (B2N 45)

She stroked the **cat**. (CHo 1898)

... the man waited, like a **cat** waiting for a mouse. (FSJ 75)

A more advanced task, though of the same type, might stimulate language awareness and (cross-)linguistic reflection by targeting words or expressions that have multiple meanings and/or lack a direct counterpart in Norwegian. One such a word is *mind*, illustrated in Table 23.2. *Mind* can be a noun or a verb, and it enters into a number of (relatively) fixed expressions in both functions. Besides, from a Norwegian point of view, *mind* is a hard word because there is no Norwegian word with exactly the same meaning; rather, we use a number of different words to express the meanings of *mind*, depending on the context. Here are some suggested activities using a concordance of *mind* (see Table 23.2):

- Identify the expressions with *mind* that the learners are familiar with. This can be followed up by, for example, finding equivalent expressions in Norwegian (or other languages that the learners speak) or synonymous expressions in English.
- Identify the expressions with *mind* that the learners are unfamiliar with. Try to work out what they mean from the context. Search for more examples to check if the assumptions are correct.
- The students may search in a corpus for more examples of the expressions in the concordance, to see if there is any variation in the wording. For example, searching for “on \* mind” from line 3 (where the \* can stand for any word), we find *his, my, her, your, the* and *its* between *on* and *mind*. *On his mind* occurs in expressions with *be* (something is on his mind) or *have*, as in line 3. Since *the* is different from the rest, it is worth exploring. *On the mind* occurs mainly in formal contexts, for example *Other ideas seem to operate on the mind with great force*. It combines with other verbs than *have* and *be*, and seems to be less of a fixed expression.
- If the corpus contains different genres, it is also possible to find out where the expression is typically used. For example, *never mind* and

*I don't mind* both occur mainly in conversation and fiction, and not so much in the other genres in the BNC.

- Focus on complex constructions with *mind*, such as the pattern in lines (4) and (9) and produce similar sentences with this pattern.

**Table 23.2** Concordance of *mind* from the British National Corpus.

1	ABX 3249	I don't <b>mind</b> .'
2	ACB 670	' <b>Mind</b> you,' Gazer continued, 'I'll remember what you said about them pill-boxes.
3	ACV 838	Mungo had so much on his <b>mind</b> that he was unable to concentrate on Mary Ann's stories.
4	ACV 2293	'Don't you <b>mind</b> him stealing your father's eggs?'
5	AEB 1717	'Give me time to make up my <b>mind</b> .
6	AT4 3311	'Well, Deirdre's a determined girl once she's set her <b>mind</b> on anything, I'll give her that.
7	B0B 2029	The thought of Joe came at once into her <b>mind</b> , but where could Joe be?
8	BMS 2486	Can't you <b>mind</b> your own business?'
9	BMU 1507	If people don't <b>mind</b> my having no degrees, I could give a few music lessons!
10	BMU 2671	Oh, never <b>mind</b> — that's beside the point.
11	C85 3475	I wouldn't <b>mind</b> taking up with her ... '
12	C87 1798	Snooker's a lot more fun when you don't have to wear a tie and waistcoat — who in their right <b>mind</b> would want to do that?
13	CE0 331	But we didn't <b>mind</b> , we soon forgave him
14	CEJ 216	He wrote with a particular audience in <b>mind</b> and therefore emphasised the points of interest most suited to that audience.
15	CFJ 265	'She's made up her <b>mind</b> .'

Another task type, which can be easily adjusted to the learners' needs and proficiency level, is to search for a word or phrase with multiple meanings and identify instances of the different meanings from the concordance. For example, the verb *appreciate* can mean "value" or "realize/understand", as illustrated by the following sentences from the BNC.

What a nice thought! I'm sure they'd **appreciate** it.

First, you must **appreciate** that a helicopter produces two different types of lift.

An area in which the corpus is unbeatable is *collocations*. A collocation is a combination of two or more words that routinely occur together. It is not necessarily an idiom, since collocations may have literal meaning and need not constitute complete phrases. But collocations represent the linguistic habits of native speakers of a language. We can find collocations in the corpus by searching for component parts of them, for example *a piece of\**, where the asterisk represents an unspecified word. The ten most frequent nouns to follow *a piece of* in the Corpus of Contemporary American English are *paper, cake, music, land, wood, bread, meat, furniture, fruit, legislation*. This information can be used in working with fixed expressions with *a piece of*, but also in a more cognitively advanced task where learners are asked to consider the meaning of *a piece of* in each collocation. They will then realize that *piece* can sometimes be translated by the Norwegian *stykke*, for example in *a piece of cake*, but not in *a piece of furniture*. Furthermore, if *a piece of cake* is used with metaphorical meaning, the Norwegian *et kakestykke* does not work as a translation. The concordance of *a piece of cake* will reveal whether the literal or the metaphorical meaning is more common.

## The internet as corpus?

The internet is by far the largest existing digital collection of texts, and it can be used in language studies in much the same way as a corpus. Besides finding information on specific topics, a search engine such as Google can also show how an expression is used (especially if we use quotation marks around the expression we search for). So if we are wondering, for example, whether it is more common to say *texted him* or *texted to him*, Google will show up more than 70 times as many hits for *texted him*. A clear advantage of Google is its easy access to huge amounts of material in most languages. It is a goldmine for information on the use of new words and expressions, which may not have found their way into dictionaries or established corpora yet. At the same time one must be critical when using the internet as corpus. For example, many English texts on the web have been written by people who are not native speakers of the language. There is thus a greater risk of finding unidiomatic expressions on the web than in a corpus. This risk can be reduced by using advanced search mode, which enables us to limit the

search to websites in countries where English is the (major) first language. Another disadvantage is that, unlike a concordance, the output of a Google search is not displayed in a way that makes it easy to see patterns of use.

There are, however, web interfaces that show search output in the form of concordances, such as WebCorp (see Figure 23.3). The concordance format makes it easier to focus on the linguistic expression, and a bit harder to be distracted by the content. WebCorp sorts the hits according to the websites they come from. This is useful because the hits may come from dictionaries, which is not the type of material one is primarily after in corpus linguistics. WebCorp gives a choice of three alternative search engines, of which FAROO and Guardian Open seem to give the most relevant results for those who are interested in studying natural language use rather than definitions and examples from dictionaries.

1) <http://www.adweek.com/digital/facebook-is-working-on-technology-that-lets-you-type-and-control-vr-devices-with-your-mind/>

Text, Wordlist, text/html, UTF8 (Content-type), 2017-01-01 (Copyright footer)

1: Lets You Type and Control VR Devices With Your **Mind** Share About / AdvertisingMedia Kit Sponsor

2: Lets You Type and Control VR Devices With Your **Mind** Could a 'brain mouse' be coming? By Marty Swant

3: is creating a way to control VR and AR with the **mind**. Getty Images Share

---

4) <http://dailycaller.com/2017/04/21/trudeau-says-trump-is-willing-to-change-his-mind-unlike-many-politicians/>

Text, Wordlist, text/html, UTF8 (Content-type), 2017-01-01 (Copyright footer)

11: Trudeau Says Trump Is Willing To Change his **Mind** 'Unlike Many Politicians'

12: ideology makes him open to changing his **mind**. "I've learned that he listens," Trudeau told

**Figure 23.3** Concordance of *mind* from WebCorp.

Alternative solutions are offered by the NOW corpus (*News-On-the-Web*) and *Querying Internet Corpora*, where existing webpages are organized as a corpus. The NOW corpus is updated every day. Table 23.3 shows some concordance lines from Querying Internet Corpora. The clickable abbreviations in the leftmost column give the sources of the examples.

**Table 23.3** Concordance of *funny* from Querying Internet Corpora

nrs	Well, my God, he' s lost it. This is n't	<b>funny</b>	! “ And then the more they read,
gou	“I just thought the word ‘goat’ was	<b>funny</b>	, “ the court then pressed her to explain
ofw	the mayor controlling the system. It's	<b>funny</b>	, Joel has a sense of grass roots
nik	Gil Cates, said that, “ He is smart, quick,	<b>funny</b>	, and loves movies.
pnq	fun and wrong, but it ended up being very	<b>funny</b>	, and the Nick/Phyllis affair.

## “Do-it-yourself” corpora

Corpus methods can be used with all kinds of texts. There are several corpus tools (concordancers) for this purpose, one of which is called AntConc. It is freely downloadable and designed to be used with files in plain text format; that is, with the extension .txt at the end of the filename. If we have a file in a different format, such as .docx or .odt, we can convert it to plain text in the word processor by means of the function “Save as”. The use of AntConc is explained on the website of its originator, Laurence Anthony (see the list of corpus resources below). Concordancers such as AntConc make it possible to use corpus methods creatively. For classroom purposes it may be interesting, for instance, to use literary texts or texts produced by students and explore them as one would a corpus.

A small study of J. D. Salinger's novel *The Catcher in the Rye* can illustrate how corpus methods can be applied to a literary text, after uploading it in AntConc in txt format. The WordList function of AntConc can provide an alphabetical list of the words in the text, but, more interestingly, it also produces a *frequency word list*, in which the most frequent words are listed first and the least frequent ones last, and which shows the number of times each word occurs. In *The Catcher in the Rye*, the most frequent word is *I*, followed by *and* and *the*. This is not unexpected, because all texts contain

a lot more so-called function words (such as pronouns, articles and prepositions) than content words (nouns, verbs, adjectives and adverbs). For literary studies we may be more interested in the content words. One of the most frequent ones in *The Catcher in the Rye* is *old*. The concordance lines reveal that *old* often appears in front of the name of a person who is objectively quite young, so in those contexts the word means “well known to the main character” rather than “aged”. Furthermore, the word *goddam* is fairly frequent, which will not surprise anyone who has read the novel. Sorting the concordance by the word to the right of *goddam* we discover that all kinds of things are characterized by this adjective. And the concordance of *pretty* shows that this word is used as an adverb (for instance in *pretty good*, *pretty nervous* and *pretty ugly*) more often than as an adjective meaning “good-looking”.

Concordances of central words in a literary text can say something about both style and content in the text. So can searches for repeated combinations of words. If we search in *The Catcher in the Rye* for repeated sequences of at least four words, we find *I don't know*, *all of a sudden*, *for God's sake* and *but I didn't* at the top of the list (AntConc counts 's and n't as words). The phrase *to know the truth* is also among the ten most frequent ones. The list of repeated sequences highlights a good number of negative sentences with *I* as subject (such as *I don't know*, *I didn't feel*, *I didn't say anything*), which may suggest a negative attitude on the part of the main character. *To know the truth* is part of the expression *if you want to know the truth*, which the main character uses often, perhaps to emphasize his truthfulness.

Using corpus tools with student texts can also be a useful exercise. Merging texts by all the students in a single text file should normally secure enough anonymity for the file to be used in class (but check whether any personal/confidential information should be removed first). Some issues that can be worth investigating include:

- What words are frequent in the students' texts compared to texts by native speakers of English on a similar topic? Are there words that are very frequent in the student texts that could be omitted or replaced by other words and expressions?

- How are words combined? For example, what words occur around *man* and *woman*?
- How do the students use particular words that they should master, for example linking adverbs (such as *however*, *in fact*) or technical terminology in their field?

It may be noted that the study of learner corpora is a research field of its own. See for example Hilde Hasselgård and Stig Johansson (2011) for an overview. In a Norwegian context, this research has deepened our insight into particular features of the language produced by Norwegian learners of English, including problems, successes, and features that set this learner group apart from other learner groups and/or native speakers of English.

## Free online corpus resources

Below follows an overview of some of the corpus resources that are available online. Some of these resources require registration, but no subscription fee. No matter which corpus we choose to work with, it is useful to remember the following advice about sensible corpus use:

### Common corpus sense

- Know your corpus! You need to know what kinds of texts the corpus contains to tell what you can use it for.
- The corpus cannot show you what is right and wrong, only what is common usage.
- Different corpora are suitable for different purposes.
- The fact that something is not found in a corpus is not proof that it does not exist in the language.
- The corpus itself cannot answer a linguistic question. The corpus user must interpret the concordances and make sense of the data.

## The corpora referred to in this chapter

- British National Corpus (BNC): Contains 100 million words of written and spoken British English, from the early 1990s. <https://www.english-corpora.org/bnc/>
- Brown Corpus: Contains a million words of written American English from the 1960s. Searchable from Sketch Engine (Open corpora) at <https://app.sketchengine.eu/#open>
- Corpus of Contemporary American English (COCA): one billion words of written and spoken American English from the 1990s up to present (the corpus is continually updated). <https://www.english-corpora.org/coca/>
- GLOWBE Corpus: A large collection of texts drawn from the internet, representing first- and second-language varieties of English from around the world. <https://www.english-corpora.org/glowbe/>

## Other corpora with search interfaces

- A range of large, English-language corpora are available from <https://www.english-corpora.org/>. Apart from the BNC and COCA, these include the NOW corpus, a corpus of American soap operas, and more.
- Some corpora are freely available from Sketch Engine (Open corpora) at <https://app.sketchengine.eu/#open>. Note particularly the British of Academic Written English Corpus, which contains student essays written at British universities. More corpora are available by subscription from the same provider.
- Michigan Corpus of Upper-Level Student Papers (MICUSP) contains highly rated American university student assignments in a range of disciplines. Available at <http://micusp.elicorpora.info/>. The interface may not work in all browsers (it works in Chrome and Firefox). The related Michigan Corpus of Academic Spoken English is found at <https://quod.lib.umich.edu/m/micase/>.

## The web as corpus and web-based corpora

- WebCorp: <http://www.webcorp.org.uk/live/>
- Querying Internet Corpora: <http://corpus.leeds.ac.uk/internet.html>

- The NOW corpus (News-On-the-Web). From 2010 until “yesterday”, updated every day. <https://www.english-corpora.org/now/>

## Databases of downloadable texts

- The Gutenberg Project: <https://www.gutenberg.org/> offers thousands of downloadable books. If in plain text, the downloaded file(s) can be used with AntConc. Recent literature is not included for copyright reasons.
- Oxford Text Archive develops, collects and preserves electronic literary and linguistic resources for use in Higher Education, in research, teaching and learning. See <https://ota.ox.ac.uk/>.

## Corpus tool (to use text files as a corpus)

AntConc is a corpus tool that can be easily installed at no cost on a pc or a mac. The tool can be used with texts in plain text (\*.txt) format. AntConc, and instructions for its use, can be found at <http://www.laurenceanthony.net/software/antconc/>.

## Learner corpora

Many of the available corpora of learner language (English and other languages) are listed at <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html>. Note that relatively few of these corpora are freely available, mainly for copyright reasons.

## YouTube

A YouTube search for “Corpus linguistics” gives a number of hits. The following presentations are recommended:

- Tony McEnery: Corpus linguistics: method, analysis, interpretation. <https://youtu.be/YJTM3i5HxsQ>
- Randi Reppen: Using corpora in the language classroom. <https://www.youtube.com/watch?v=Qf46lOnMCfs>
- Michaela Mahlberg on corpus linguistics and literature: <https://www.youtube.com/watch?v=kvbrp5PqNxw>

## Corpus exercises for secondary school students

Dypedahl, Magne and Hilde Hasselgård. 2006. *Exploring English*. Website. <http://exploringenglish.cappelendamm.no/c302684/artikkel/vis.html?tid=329736>.

## Concluding remarks

The linguist M.A.K. Halliday (1991) argues that “the immense scope of a modern corpus, and the range of computing resources that are available for exploiting it, make up a powerful force for deepening our awareness and understanding of language” (p. 41). Both teachers and learners of English can benefit greatly from learning how to use a corpus: they will have practically unlimited sources of information on English language use at their disposal. The corpus can offer invaluable assistance where our intuitions (even in our first language) are insufficient, for example when it comes to studying collocations, making sense of ambiguous words and expressions, and assessing how frequent an expression is compared to another. In fact, research has shown that people’s intuitions about frequency are unreliable because we notice what stands out rather than what seems normal. But in order to benefit from the corpus, we must know how to use it. That means knowing how to ask good questions, how to find appropriate ways of searching in the corpus and, most importantly, how to interpret the search results. Anyone who has acquired these skills will soon discover that corpus use is addictive. And unlike many other addictions, corpus linguistics is purely beneficial.

### Reflection questions

1. Do you agree that corpus work represents a descriptive approach to language and language learning? What are the reasons for your agreement or disagreement?
2. What do you see as the benefits and problems of using authentic (corpus) material in English language teaching?

3. Think of ways in which learners can be guided into making discoveries about language, with or without the aid of corpora.
4. Discuss whether (and/or how) you could use corpus methods with literary texts as a bridge between active language learning and literary interpretation.
5. How would you design corpus tasks that are relevant to particular groups of learners, taking account of age, proficiency level and so-called "specific purposes" (for example English in vocational training)?

## References

- Biber, D., Conrad, S., & Leech, G. (2002). *Longman student grammar of spoken and written English*. Longman.
- Halliday, M.A.K. (1991). Corpus studies and probabilistic grammar. In K. Aijmer, & B. Altenberg (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik* (pp. 30–82). Longman.
- Hasselgård, H., & Johansson, S. (2011). Learner corpora and contrastive interlanguage analysis. In F. Meunier, S. De Cock, G. Gilquin, & M. Paquot (Eds.), *A taste for corpora. In honour of Sylviane Granger* (pp. 33–82). Benjamins.
- Hidalgo, E., Quereda, L., & Santana, J. (Eds.). (2007). *Corpora in the foreign language classroom*. Rodopi.
- Johansson, S. (2011). A multilingual outlook of corpora studies. In V. Viana, S. Zyngier, & G. Barnbrook (Eds.), *Perspectives on corpus linguistics*. Benjamins.
- Mukherjee, J. (2002). *Korpuslinguistik und Englischunterricht. Eine Einführung*. Peter Lang.
- O’Keeffe, A., McCarthy, M. & Carter. R. (2007). *From corpus to classroom. Language use and language teaching*. Cambridge University Press.
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31, 205–25.
- Sinclair, J. M. (Ed.) (2004). *How to use corpora in language teaching*. Benjamins.