The Use and Abuse of Meaning Technologies

Philip Hubbard
Stanford University, USA

Invited paper: Ontario TESL Conference, Toronto, Canada, November 2000
To appear in *Contact*: Autumn 2001.

In simple terms, meaning technologies are features of language learning materials that are based on technology and make the meaning of a language item accessible to a learner. This paper begins by discussing the general relationship of meaning and meaning aids to language learning. It then introduces the concept of meaning technologies and explains why they are becoming increasingly important in computer-assisted language learning. It concludes by offering guidelines for using these technologies appropriately.

We can begin by asking how learners extract meaning from language forms, including live conversations, lectures, audio or video recordings, and written texts. We know that like native speakers, language learners use their linguistic knowledge for this. They then integrate context and background knowledge, and use inferential strategies to link that knowledge to the language forms. Given their lack of familiarity with many of the forms they encounter, learners also use *meaning aids* when they are available and circumstances allow it.

Meaning aids are tools that serve to make the meaning of language items accessible. Traditionally, they include dictionaries and glossaries, which help make the meaning of words clear; summaries, paraphrases, and explanations, which assist in the interpretation of larger texts; audio and text translations, which link the target language back to the native equivalents; and scripts for spoken forms, which make their meaning more accessible to those whose reading proficiency outstrips their listening proficiency.

It is generally accepted that meaning aids can improve the comprehension of items within a specific text and therefore the text as a whole. However, research does not tell us much directly about how to use meaning aids effectively to improve language *learning*. In other words, we do not have a clear answer to the key question: How can a learner use a meaning aid on one text in a way that will improve comprehension of some subsequent one?

Still, research can provide us with some important generalizations to guide practice in this area. SLA researchers appear to be in general agreement that comprehension is a necessary, though probably not sufficient, condition for acquisition (Long, 1990). Any view of comprehension-driven second language acquisition, whether it is primarily unconscious (Krashen, 1985), or acknowledges the need for conscious attention and awareness (Schmidt, 1995) relies on not just comprehending meaning, but also somehow linking that meaning to an unacquired L2 rule or form. It appears to be the case, however, that normal language processing involves *automatic* linking of form to meaning, with the original form then lost and only elements of the semantic content retained (Clark & Clark, 1977). Further, psycholinguistic research on cognitive capacity limitations cited by

Skehan (1998) suggests that multi-channel input (such as text and sound simultaneously) may actually put a greater burden on a learner's limited processing capacity than a single channel would, particularly if the material itself is cognitively challenging.

Additionally, since a certain rate of processing appears necessary to maintain text coherence, the use of meaning aids, such as dictionaries, can interrupt the flow of sentence and discourse comprehension. Recent studies using MRI (magnetic resonance imaging) are adding support to the theory of the brain being highly modularized with systems that collaborate in processing language in parallel but may be in competition for processing resources when dealing with material that has not become automatized (see Grodzinsky et al. (2000) for a collection of relevant articles). In particular, work by Smith & Geva (2000), though not directly connected to acquisition, shows that the traces of phonological processing necessary for creating a form-meaning link for new items last only a few seconds in working memory. It is thus common to hear a new form and lose its memory trace almost immediately.

The results cited from these various research disciplines are somewhat tentative and open to other interpretations. However, taken together they are consistent with the following assumptions which inform the discussion in the remainder of the paper:

- Meaning aids can allow learners to access meanings of unfamiliar language items, making input comprehensible.
- Meaning aids can get in the way of overall comprehension by interrupting the flow of processing.
- Meaning aids can disable the desired form-meaning connection by causing the learner to focus on a form that is different from the original.

The preceding can be summed up in a statement that most ESL teachers already know from experience: meaning aids can be either useful or not for the learner depending on how and when they are used.

Having established that meaning aids represent a two-edged sword, let us now turn back to the topic of meaning technologies. *Meaning technologies* are meaning aids that use technology to help make the meaning of a language item accessible on demand. For the most part they are representations of the more traditional meaning aids, but they can typically be accessed instantaneously through the click of a button, increasing their convenience and efficiency significantly. We can recognize two types of meaning technologies: dedicated and automated.

Dedicated meaning technologies are those which are deliberately programmed into text, audio, or video material. They include such features as closed captioning in videos and multi-media computer programs, online transcripts for audio and video materials on the web, hypertext glossaries and explanations, translations (including multilingual soundtracks for DVDs), and slower or otherwise simplified audio options. Besides the closed captioning for videotapes, these dedicated meaning technologies can be found in large numbers of CD-ROMs for language learning, and they are appearing increasingly in programs on the web (see the appendix for web-based examples).

Automated meaning technologies are general applications which can be used directly by software developers and language learners without additional programming. This further increases both their power and their appeal. As they are less familiar to many than the dedicated technologies described above, I will present them in a bit more detail. I will focus on four of these, though there are others which are likely to appear in the not-too-distant future.

Hyper-linked dictionaries are applications that automatically link the words in a text to one or more dictionary entries, similar to the way word processor-dictionary or word processor-thesaurus combinations allow writers to highlight a word and then do a lookup automatically. A promising commercial application in this area is Sentius Corporation's RichLink (see appendix), which is a web-based technology that allows a single click on a word to bring up entries in linked dictionaries or other data sources. The entries appear in a small pop-up box instantaneously so text coherence is less affected. The dictionaries can be monolingual (they have licensed the *American Heritage Dictionary*, for example) or bilingual. While the pages themselves must be processed by the Sentius servers, once processed they can be read on the web by anyone with a recent version of Netscape or Internet Explorer. Two currently free applications in this category are Voycabulary and Babylon (see appendix). Voycabulary works by allowing users to type in a URL from the Voycabulary site which then indexes all words there for lookup by clicking on them. Babylon is the most powerful of these applications in that it allows users to instantly look up words in web pages, electronic texts, emails, and word processing documents without separately processing them, and it is also available offline.

Machine translation applications represent a dream that is still far from practical realization if the goal is something approaching native-like output. However, machine translation sites do exist on the web, and it is easy to demonstrate that they can still be useful in helping learners comprehend the gist of advanced texts. The most common translation engine on the web is Systran, which was developed by Xerox. The Alis and Systran sites (see appendix) are good places to check translations from English into other languages and vice versa. All that is necessary is to copy and paste text into their translation box and click the "translate" button.

Auto-summarizers are a little-known meaning technology in language teaching circles, but they are used increasingly in the business world. These applications scan a text and pull out phrases or sentences based on routines that look for keywords and tag locations most likely to contain information about the main ideas of a text. They can be set to a specified limit (e.g., 10 sentences) or a percentage of the whole (e.g., 10%). This is actually embedded as a feature in recent versions of Microsoft Word. To see if it is installed, go to the Tools menu and look for AutoSummarize.

Text-to-speech applications represent the final automated meaning technology to be discussed. Just as captions can aid more literate learners in accessing audio texts, text-to-speech programs can aid those whose spoken language skills are stronger in developing reading proficiency by speaking the words out loud. While the synthesized voices typical

of text-to-speech applications remain noticeably artificial at best, they can still aid some students in their comprehension of the printed word, particularly in light of the vagaries of English spelling. Lucent Technologies has a text-to-speech demonstration site and EnglishPractice.com has a site that uses an animated character (based on Microsoft Agent) to read the texts out loud (see appendix). Additionally, most speech recognition software and some versions of the Macintosh operating system have text-to-speech features.

The uses of meaning technologies are similar to those for the more traditional meaning aids. The following are a few examples.

- Closed captioning can be used when watching a video for the first time to make more of the material comprehensible and then turned off for a second listening to push a focus on the spoken forms. Alternatively, it can be off during a first viewing to determine what can be comprehended naturally and then turned on for a second viewing to help the learner become aware of unknown forms.
- Hypertext links can be used during a close reading to look up a word meaning, though ideally after attempting to determine it from the context.
- Translations of individual words, sentences, or even whole texts may help learners access material that would otherwise be beyond their reach, particularly if they are working with authentic documents and their proficiency is low. It is also possible to take a "bad" machine translation from the students' native language into the target language to see if they can determine the mistakes and infer their source.
- Text-to-speech can be used to support literacy development, and can even be turned around to assist more proficient readers with the pronunciation of familiar words they have only read and not heard.

The key to using meaning technologies appropriately is to understand clearly what the objective of reading or listening to the text is. If it is only comprehension to serve some other purpose, then anything which makes the meaning accessible is good, including the free use of accompanying meaning technologies. If, however, the goal is to improve language proficiency, then one or more of the following should be considered.

1. Use the meaning technology as part of a pre-reading or pre-listening activity. After getting the meaning, return to the original form and read (if printed text) or listen (if audio or video) without the meaning technology's support.
2. Read or listen for the first time without any meaning technology, processing as much as possible of the target text. Then, access the meaning technology selectively in the second pass.
3. Use the meaning technology for the first part of the text, e.g., turn on closed captioning for the first half of a video segment or allow lookup of hypertext-linked words in the first half of a reading. Then read or listen without the technology's support for the second part.
4. If there are grammatical structures, phrases, or words to be learned, try to capture the direct form-meaning relationship for focused attention, rehearsal and review. For example, get learners to make a list of some or all of the words looked up with a hypertext glossary for a personal dictionary.

The abuses of meaning technologies occur when learners use them in a manner that subverts the objectives of the task or undermines the long-term language learning goals of the curriculum (and presumably of the learners themselves). For example, learners who are significantly better at reading than listening may overuse captioning in videos and software. Cognitive capacity theory suggests that the processing of the printed word, particularly if it lags behind the spoken word, may replace the native spoken forms in working memory. Learners who have translations readily available to them may continue to rely on them at the expense of building automaticity in going directly from form to meaning in the target language. Similarly, learners who feel the need to look up unfamiliar words rather than develop strategies for inferring their meaning from context will find the ease of using hypertext glossaries and dictionaries an even stronger temptation.

Finally, it is unfortunately the case that not all learners understand or care about the learning objectives of an activity, especially when the language course is a requirement and motivation is therefore extrinsic. In these situations, meaning technologies may represent appealing shortcuts to the completion of in-class exercises and homework.

Clearly, there is a need for more research in the use of meaning aids in general and of meaning technologies in particular. Realistically, however, we can not depend on such research to appear anytime soon in a form comprehensive enough to determine practice. This is especially true given that the structure and power of meaning technologies changes so rapidly. However even without direct guidance from research, there are several steps that software developers and language teachers can take to support appropriate uses and limit abuses.

Developers need to abandon the view that dedicated meaning technologies are automatically useful additions to any piece of software and every activity within that piece. To the extent that tutorial software embodies good teaching, learners should be encouraged by the structure of the program to utilize meaning technologies in ways similar to those an experienced human tutor would promote. Similarly, teachers need to be aware of the meaning technologies available to learners and design both in-class and homework activities that are not easily and obviously done more quickly but less effectively by using these aids.

Most importantly, both developers and teachers need to train learners in appropriate uses of meaning technologies. Software with dedicated meaning technologies should have tutorials designed to give both examples and explanations of learning strategies using those features. Teachers should spend time not only introducing effective uses of the features but also observing learner use and retraining as needed. If learners are to be given the kinds of power that meaning technologies afford, then they should also be given enough understanding of learning processes to be able to use them in an informed manner.

Meaning technologies are still in their infancy, but there is a tremendous amount of time and money currently being spent on developing some of these for large-scale commercial purposes. We can expect the technologies mentioned here to become more accurate and less expensive in the next few years. We can also expect some of the current dedicated technologies, such as closed captioning, to become automated as more reliable speech-to-text dictation systems are devised. It may be common a decade or two from now for a learner to query an automated teaching agent with the question "What does that mean?" and get a an instant, useful, and accurate response. We need to begin now to learn how to control these technologies so that the answer to that question leads not just to meaning but to learning.

References

Clark, Herbert & Clark, Eve (1977): *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

Grodzinsky, Yosef, Shapiro, Lew, & Swinney, David (eds.) (2000). *Language and the Brain.* New York: Academic Press.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Kurzweil, Ray (1999). *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Viking.

Long, Michael (1990). The least a second language acquisition theory needs to explain. *TESOL Quarterly* 24.4, 649-666.

Schmidt, Richard (ed.) 1995. *Attention & Awareness in Foreign Language Learning.* Honolulu: University of Hawaii Press.

Skehan, Peter. 1998. *A Cognitive Approach to Language Learning* New York: Oxford University Press.

Smith, Edward & Geva, Anat. 2000. Verbal working memory and its connections to language processing. In Grodzinsky, Yosef, et al. (ed.). *Language and the Brain.* New York: Academic Press.

Appendix: Links for Examples of Meaning Technologies

DEDICATED MEANING TECHNOLOGIES ON THE WEB (checked as of 3/11/01)

http://www.englishbaby.com/
EnglishBaby includes weekly dialogues with characters in a college campus setting, with sound, transcripts, and a hyperlinked glossary of key words and idioms. There is an archive of previous weeks, and it is free at the moment.

http://www.englishlistening.com/
This commercial site has a listening section with transcripts and comprehension questions. There are some free samples, but after that it costs.

http://www.pbs.org/newshour/media
This has the video and transcript of the daily PBS program "The News Hour with Jim Lehrer." It works best for learners with advanced listening skills and an interest and background in the stories the show focuses on.

http://www.learnenglish.org.uk/
This is a British Council site with a lot of useful material. Learners can click on "stories" to get both text and audio of stories online.

http://www.trstone.com/
This site has free online demos of Fairfield's The Rosetta Stone, language learning software for English and many other languages with a number of dedicated meaning technologies. In particular, it has examples of using slowed versions of example sentences.

AUTOMATED MEANING TECHNOLOGIES ON THE WEB

http://www.lucent.com/speech/welcome.html
The site of Lucent Technologies text-to-speech system demonstration.

http://www.englishpractice.com/
This has text-to-speech on an ESL site. From the home page click "Study by Course" > "TOEFL Course" > "Read to me". A character (Microsoft Agent's "Merlin") will read whatever you type or paste into the text box.

http://www.richlink.com/
This site has samples of Sentius Corporation's RichLink technology, a commercial system for embedding dictionaries and other databases directly into web pages for hypertext linking.

http://www.voycabulary.com/
This site allows users to type in URLs and then processes the requested page to provide automatic hypertext links.

http://www.babylon.com/
Babylon is a very popular and powerful hypertext linker. After a short download of a resident application, it is possible to click on words in almost any text format and get a pop-up definition. Babylon has both a free ad-supported and inexpensive ad-free version.

http://www.alis.com/
Alis presents their version "gist-in-time" machine translation using Systran.

http://www.systransoft.com/
This is the homepage for the web version of Systran, the most widely used of the machine translation applications.

http://cnniw.yellowbrix.com/
This is CNN's industry watch page—a good example of using autosummaries. You can view the summaries and the whole articles they're attached to. CNN uses Inxight's Summary Server http://www.inxight.com/

http://www.copernic.com/products/summarizer/
This is an example of a consumer-level autosummarizer with more functionality than the one in MS-Word.