CHAPTER

13

Language and Technology

SECTION 13.1: CELL PHONES

SECTION 13.2: TEXTING

SECTION 13.3: NATURAL LANGUAGE PROCESSING

SECTION 13.4: ARTIFICIAL INTELLIGENCE

For most of our existence as humans, speech was the only mode of language available to us. We lived in small bands as hunter-gatherers, and speech gave us the ability to coordinate our endeavors in ways that our primate ancestors could not.

The invention of agriculture within the last 10,000 years sparked a technological revolution. Our numbers increased, and we started organizing ourselves in city-states. Kings needed to keep track of their tax revenues, and businessmen their transactions.

Writing evolved out of the tally sheets of early accountants. But more than just being a memory aid, writing became a new mode of communication. For the first time, humans could communicate across both distance and time. However, the writing mode also forced us to use language in a more careful manner to compensate for the lack of facial expressions and vocal inflections. Even today, new technology can influence the characteristics of the language we use.

The invention of writing enabled us to exchange messages with others at great distance, but those messages had to be physically carried from sender to receiver. That changed in the middle of the nineteenth century with the invention of the telegraph. Now written messages—transformed into a series of dots and dashes—could be sent rapidly along wires across great distances.

Sending telegrams was expensive, though, and a “telegraphic” manner of writing evolved that used only content words. This style was also adopted for newspaper headlines and postcards. A few decades later, the invention of the telephone enabled conversation at a distance for the first time. Still, a number of linguistic conventions evolved as we modified our natural form of communication to the new technology.

During the last few decades, many new modes of communication have emerged, from email to Skype, from SMS to Twitter. Each of these media has particular restrictions, and we quickly adapt our language use to fit them. In this chapter we explore how new technologies impact the way we use language.

SECTION 13.1: CELL PHONES

- Driving is significantly impaired by talking on a cell phone, and distracted driving leads to thousands of deaths and hundreds of thousands of injuries each year; talking on a cell phone impacts driving even more than being legally drunk.
- Hands-free cell phones are no improvement, since it is not the manipulation of the device that causes distracted driving; rather, it is the conversation itself, over a mobile device, that impairs driving. This is in contrast to listening to the radio, which does not negatively impact driving ability.
- When we multitask, we alternate our attention between two or more cognitive tasks; conversing on a cell phone draws heavily on attentional and cognitive resources, and drivers have difficulty diverting those resources back to the driving task when road conditions are hazardous.
- Conversation is an activity that engages joint attention. In the case of cell phone conversations, joint attention is drawn away from the driving environment; however, in the case of conversing with a front-seat passenger, joint attention is often directed toward road conditions. As a result, drivers have fewer accidents when driving with a passenger than when driving alone.
- Cell phone conversations lead to inattentional blindness, which is the failure to see or to remember objects or events directly in our visual field because our attention has been drawn elsewhere; both drivers and pedestrians are less likely to notice potential hazards when using a mobile device.
- Although cell phone calls in public are often perceived as obnoxious because they are loud, experimental studies indicate that only being able to hear half of the conversation is what actually makes them annoying.

Cell phones have truly turned the world into a global village. For the first time in the history of our species, a human being can initiate a conversation with another person without even knowing where on the planet he or she may be located. In fact, "Where are you?" has become a common way to start a cell phone conversation.

More than 80% of American adults own cell phones, and the percentage is similar throughout the developed and developing world (Beeson, Higginson, & Rising, 2013). Among young Americans in the eighteen- to twenty-nine-year-old range, ownership is virtually universal, at around 96% (Morrill, Jones, & Vaterlaus, 2013). For this population, cell phones aren't just a convenient tool for communication. They're an integral part of a young person's social identity, as brand name and model provide important information about the personality of the owner.

Many adolescents and young adults develop strong attachments to their cell phones (Weller et al., 2013). It's not unusual for young people to experience separation anxiety when they lose or misplace their mobile devices. In surveys, respondents in this age group often report that they feel naked without their cell phone and that they'd rather lose their wallet than their cell phone. The need to communicate is essential for well-being, and cell phones have become a major mode of communication for young adults.

Risky Business

As cell phone ownership has increased, so has the use of mobile devices while driving, with a resulting rise in the incidence of traffic accidents due to distracted driving (Ferdinand & Menachemi, 2014). A quarter of all accidents and fatalities on American roads are caused by drivers who are distracted by their mobile devices (Sanbonmatsu et al., 2013). This translates to 2,600 deaths and 330,000 injuries a year, all as a result of using a cell phone while driving (Wu & Weseley, 2013).

Talking on a cell phone clearly impairs driving performance. This is true for various age groups and as measured by a number of different methods (Drews, et al., 2008). Drivers using a cell phone have slower reaction times, fail to respond to objects and events in plain sight, are more likely to miss a turn, and check their mirrors less often compared with non–cell phone users (Rivardo, Pacella, & Klein, 2008). The amount of brain activity that drivers devote to the driving task is reduced by more than a third when talking on a cell phone (Sinsky & Beasley, 2013). Driving performance while talking on a cell phone is even worse than is driving while legally drunk (Hyman et al., 2010).

Despite the known risks of using a mobile device while driving, only 3% of respondents on surveys claim that they never use their phones while driving (Park et al., 2013). Furthermore, about half of survey respondents who view cell phone driving as an "extremely serious risk" also reported that they'd used their phones while driving within the last month. Such findings suggest that drivers seriously underestimate the very real dangers of talking on a cell phone.

While hands-free devices enable drivers to keep both hands on the steering wheel, they do little to mitigate the risks of talking on a cell phone when driving (Sawyer & Hancock, 2012). This finding indicates that it's not handling the device that leads to distracted driving. Rather, driving impairment is due to the cognitive demands of conversing over a cell phone. However, recent studies have found that people who use hand-held devices while driving also tend to engage in other risky behaviors like not wearing their seatbelt (Ferdinand & Menachemi, 2014). As a result, they're not only more likely to be involved in an accident, they're also more likely to be severely injured when they do have a crash.

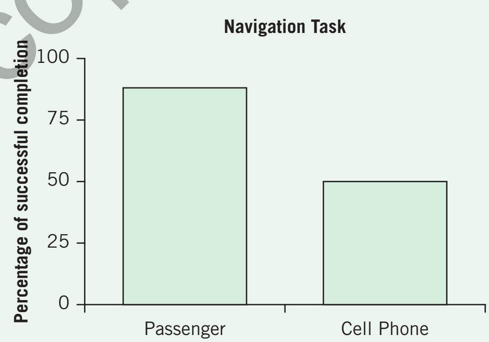
The considerable impairment caused by cell phone use while driving is counterintuitive. After all, people have been driving cars for over a century, conversing with passengers without any apparent danger. In fact, having an adult passenger in the front seat lessens the risk of an accident (Ferdinand & Menachemi, 2014). This is especially true for novice and older drivers, who are at greater risk of accident in comparison with seasoned adult drivers. Furthermore, listening to the radio or other audio media had no effect on driving performance (Strayer & Drews, 2007).

There are several reasons why conversing with a passenger isn't distracting in the way that talking on a cell phone is. First, passengers often help drivers by pointing out traffic obstacles and landmarks. Second, adult passengers regulate the flow of conversation depending on traffic conditions. In other words, when drivers and passengers converse, they tend to engage in the joint activity of shared attention to road conditions and of navigating the car to their destination (Drews et al., 2008). Cell phone conversants, on the other hand, can't do this because they're unaware of road conditions (Hyman et al., 2010).

Many studies investigating the impact of cell phone use on driving are conducted in high-fidelity driving simulators. Not only are studies easier to control in simulators, they're also safer. In one study, drivers were asked to navigate a multilane freeway for a distance of about eight miles and then to exit the highway at a designated rest stop (Strayer & Drews, 2007). During the simulated trip, the driver conversed with a friend in the passenger seat or else via cell phone. Drivers conversing with a passenger successfully navigated to the destination 88% of the time, compared with only 50% for those conversing over a cell phone.

Figure 13.1 Navigation Task

In a driving simulator task, participants were instructed to navigate to a rest area about eight miles up the road. While 88% of drivers conversing with a passenger successfully navigated to the destination, only 50% of those talking on a cell phone successfully completed the task.



Source: Adapted from data in Strayer and Drews (2007).

Some researchers have questioned whether simulator studies accurately portray the performance decrements created by cell phone use (Hyman et al., 2010). The simulated environment may in fact make driving more difficult, as can the fact that both the car and the phone are unfamiliar to the driver. Furthermore, these studies don't take into account individual differences, in that some people may be more practiced at multitasking than others. Nevertheless, it's clear that cell phone use does have some impact on driving.

Blinded by the Phone

Attention is a bottleneck in cognitive processing, and strictly speaking it's impossible to do two things at once if both of those tasks require attentional resources. Thus, **multitasking** involves *the alternation of attention between two or more cognitive tasks that take in different inputs, engage in different processes, and produce different outputs* (Sanbonmatsu et al., 2013). Anyone can chew gum and walk, because chewing is a fully automated task and walking on even terrain requires only minimal attention. Much of the time, people can walk and talk at the same time as well. Maintaining a conversation is cognitively demanding, but there's still enough attentional reserve for walking, which is largely automatic under normal circumstances. However, people do tend to stop talking when they have to navigate difficult terrain.

Talking with a passenger while driving is similar. Drivers and passengers converse when road conditions are clear, but they stop chatting during congested or hazardous

Figure 13.2 Participant Talking on a Cell Phone While Driving in a Simulator

A participant conversing on a hands-free cell phone while driving in a high-fidelity driving simulator.

Source: Drews, Pasupathi, and Strayer (2008).

conditions. Talking on a cell phone is even more cognitively demanding than conversing in person. Not all frequencies of the voice are carried over the cell phone, and noise is often mixed in with the signal. This means that we have to work harder to understand the other person, and so more attention is shifted to the task. If you're doing nothing else, talking on a cell phone isn't that challenging. But if you engage in another cognitively demanding task such as driving, you may not have enough cognitive capacity to perform both tasks well. In that case, you need to focus on one of the tasks and ignore the other. For safety's sake you should ignore the cell phone conversation and focus on the driving, but in reality that's rarely what we do.

People vary in their ability to multitask, but more importantly, perceived ability to multitask is an especially poor indicator of actual ability (Sanbonmatsu et al., 2013). In general, people overestimate their ability to focus on two tasks at the same time, and those with the most inflated view of their capacity for multitasking tend also to be the ones who are the worst at it. As a result, those drivers who habitually talk on the cell phone while driving are also the ones that are least able to regulate their attention. Researchers have linked cell phone use while driving to personality traits like impulsivity and sensation seeking. Even more dangerous is the fact that these people are oblivious to their poor performance on the road (Hyman et al., 2010).

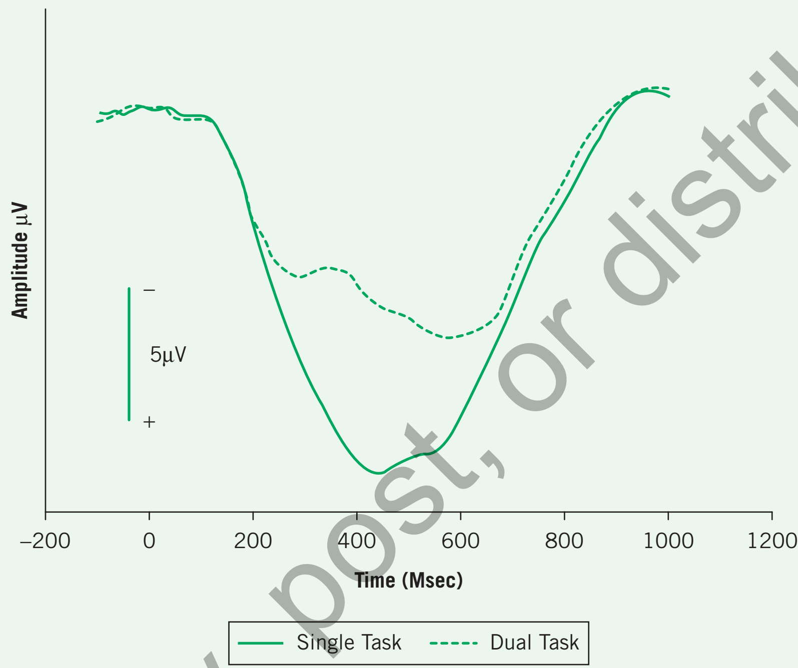
Perhaps it's because talking seems so automatic that we don't realize just how cognitively demanding it is. At any rate, talking on a cell phone often leads to **inattentional blindness**, which is *the failure to perceive or remember a stimulus directly in the visual field because another cognitively demanding task is being performed* (Hyman et al., 2010). Strayer and Drews (2007) used an **incidental-recognition-memory paradigm** to measure participant's recall of road signs they had passed while driving in a simulator. This is *a procedure in which participants are tested on their ability to remember items that they were not specifically instructed to pay attention to*.

Those participants who had driven while conversing on a hands-free cell phone only recalled half as many roadway signs as those who'd driven without talking on the cell phone. Furthermore, the traffic relevance of the roadway signs was no indicator of how likely the cell phone drivers were to remember them. That is, drivers distracted by their cell phone were just as likely to miss a hazard warning as they were an "adopt-a-highway" sign. These results show that drivers are not able to strategically allocate attention to road conditions when needed.

This inattentional blindness hypothesis is further supported by electrophysiological studies. The **P300** is *an ERP component that indicates how much attention is allocated to a task* (Strayer & Drews, 2007). Furthermore, the amplitude of the P300 is a good predictor of memory performance on a subsequent incidental-recognition-memory task. There was a 50% reduction in the amplitude of the P300 when drivers in a simulator talked on a cell phone. This finding suggests that even when they fixated on an object in the visual field, they failed to encode it as well as did drivers who weren't distracted by the cell phone. Simply put, they didn't see things that were right before their eyes because they were too busy talking.

Figure 13.3 Cell Phone Driving and the P300

Average event-related potential (ERP) elicited by the participants' perception of a brake light on the car they were following. The amplitude of the P300 component, which reflects memory encoding, was reduced by half with the participant was talking on a cell phone.



Source: Strayer and Drews (2007).

Send in the Clowns

People generally understand the potentially deadly risks of not paying attention while driving even if they underestimate their ability to regulate their attention while multitasking. The risks involved in using a cell phone while walking are less obvious, but they can be just as deadly.

Observational studies have found that people who talk on a cell phone while walking engage in the same precautionary behaviors as other pedestrians (Lopresti-Goodman, Rivera, & Dressel, 2012). Cell phone walkers visually scan their pathway for obstacles, and they look left and right before stepping out into the street. Still, they get involved in more accidents with oncoming traffic and experience more near misses than do pedestrians not using mobile devices. They also cross traffic with less time to spare and miss more safe opportunities for crossing compared with those who aren't using a cell phone. In other words, pedestrians who use cell phones appear to be experiencing inattentional blindness, in that they go through the motions of safe walking behavior but fail to detect many of the dangers they're looking out for.

Even when someone is walking in a pedestrian-only area where there's no danger from vehicular traffic, talking on a cell phone leads to impaired performance. In a set of observational studies, Hyman and colleagues investigated the effect of multitasking on walking (Hyman et al., 2010). In the first study, observers categorized pedestrians crossing an open square on the Western Washington University campus, and they also measured the time it took each pedestrian to cross the square. They found that people who were talking on a cell phone walked slower, changed direction more often, and had more near misses with other walkers compared with people who were listening to a music player, using no electronic device, or walking with a partner.

In the second study, the researchers tested the hypothesis that talking on a cell phone led to inattentional blindness while walking (Hyman et al., 2010). They hired a clown to ride a unicycle around a large sculpture in the middle of the square. They also posted interviewers to approach pedestrians as they left the square to ask them if they'd seen the clown. Only a quarter of those using a cell phone said they'd noticed the clown, while more than half of those walking alone

Figure 13.4 Photo of the Unicycling Clown

Clown unicycling around the Sky-Viewing Sculpture on Red Square on the Western Washington University campus.

Source: Hyman et al. (2010, Fig. 2). © John Wiley & Sons, Ltd.

with no electronic device saw the clown. Interestingly, those listening to an MP3 player were just as likely to have seen the clown as those who weren't, which parallels the finding that listening to the radio doesn't impair driving while talking on a cell phone does. Nearly three-quarters of those walking in pairs saw the clown, which is similar to the finding that conversing with a front-seat passenger improves driving performance. This result also points to the role of joint attention in conversation. Just as a passenger may point out a safety hazard to the driver, pedestrians who notice a unicycling clown are likely to point it out to their partner.

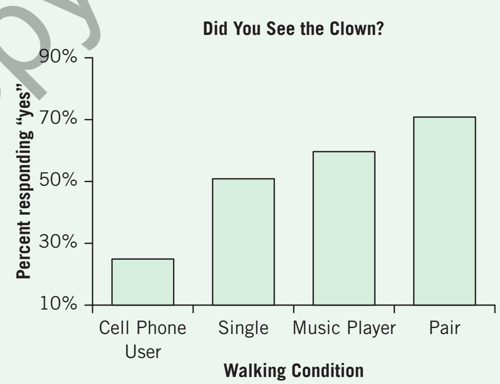
Need to Listen

Most people consider cell phone conversations conducted in public places to be annoying (Monk, Fellas, & Ley, 2004). A common perception is that people talk louder on cell phones than they do in face-to-face interactions. While this may be true, controlled studies have found overheard cell phone conversations to be rated more annoying than overhead face-to-face interactions, even when the sound volume was the same in both cases.

This has led researchers to consider *the proposal that only being able to hear half of the conversation is what makes public cell phone calls so annoying*. Monk and colleagues

Figure 13.5 Did You See the Clown?

Only a quarter of pedestrians talking on a cell phone noticed the clown on the unicycle, whereas half of those without a cell phone saw him.



Source: Based on data from Hyman et al. (2010).

(2004) tested the **need-to-listen hypothesis** by staging conversations on trains in England. Two actors sat behind a passenger and performed a scripted conversation. In the first condition, both actors spoke in a clearly audible voice, while in the second condition only one of the actors spoke in an audible voice. In the third condition, one of the actors performed the scripted conversation via cell phone with an actor in another train car.

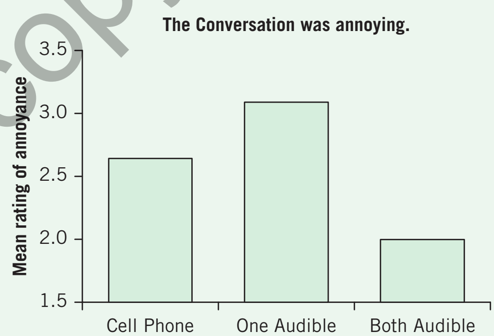
After the one-minute conversation was over, the actors left the train car. A researcher then approached the targeted passenger and interviewed him or her about the overheard conversation. On average, the respondents rated the conversation in which both partners could be overheard to be less annoying than either the cell phone or one-audible conversations, which were both rated as annoying to the same degree. Since all conversations were controlled for sound volume, this result indicates that what was annoying was the inability to hear half of the conversation.

In Sum

Taking on a cell phone impacts driving ability even more so than being legally drunk, and it leads to thousands of deaths and hundreds of thousands of injuries each year. Since holding the cell phone is not what impairs driving performance, hands-free devices yield no improvement in safety. When we multitask, we

Figure 13.6 The Conversation Was Annoying

Although all three conversation conditions were performed with the same sound volume, the condition in which both sides of the conversation were audible was rated less annoying than the two conditions in which only one side of the conversation was audible.



Source: Based on data from Monk, Fellas, and Ley (2004).

alternate our attention between two or more cognitive tasks. Talking on a cell phone makes considerable demands on cognitive and attentional resources, and we often fail to divert those resources back to the driving task when road conditions become hazardous. This isn't the case with listening to the radio, since we can easily disengage from it when traffic conditions require our full attention. Likewise, having an adult passenger in the front seat decreases the risk of an accident, since the topic of conversation often directs joint attention to the navigation task. Cell phone conversations while driving or walking are dangerous because of inattentional blindness. Because our attention is directed toward the conversation, we fail to process important objects or events in our visual field. Most people find public cell phone calls to be annoying. While loudness may account for some of the irritation, mostly these calls are obnoxious because we can only hear half of the conversation.

Review Questions

1. Explain why conversing with a passenger is safer than driving alone, and why conversing on a cell phone is more dangerous.
2. Multitasking involves engaging in two different tasks at the same time, but can you really pay attention to two different things simultaneously? Explain.
3. What is inattentional blindness? How is it measured using the incidental-recognition-memory paradigm?
4. Describe the method and results of the clown on a unicycle study (Hyman et al., 2010).
5. Describe method and results of the experiment that tested the need-to-listen hypothesis (Monk et al., 2004).

Thought Questions

1. Conduct a survey of your classmates or friends regarding attitudes toward cell phone use while driving. Ask your research participants to respond to the question: *How dangerous is it to use a cell phone? (1) Not very dangerous, (2) Somewhat dangerous, or (3) Very dangerous.* Then, follow up with the question: *How often do you use a cell phone while driving? (1) Rarely, (2) Sometimes, or (3) Frequently.* Do you detect any patterns in your data? What do they suggest?
2. Brainstorming with your classmates, generate ideas for convincing people not to use cell phones while driving. Based on your general knowledge of human behavior, which approaches do you think will be most effective?

Google It! Did You See the Clown?

There are plenty of quality videos on YouTube that demonstrate **inattentional blindness**. The **moonwalking bear** and the **gorilla on the court** are classic demonstrations. Check them out. You can also find a video demonstration of the **clown on a unicycle experiment**. For a humorous take on dealing with annoying people who talk on their cell phone in public, check out **cell phone crashing**.

SECTION 13.2: TEXTING

- Keyboards and keypads provide a means for entering text into a machine; the various keyboard and keypad layouts in use today are the result of both practical considerations and historical legacy. Human factors is an interdisciplinary field that studies how humans interact with machines.
- Statistical properties of language have been an important factor in the design of text-input systems for a century and a half. The QWERTY keyboard separates commonly co-occurring letters to prevent mechanical jamming, thus allowing for rapid typing; predictive text software enables one-thumb text input at an efficient speed.
- Texting on a cell phone is an inexpensive and private means of communication that meets the needs of adolescents and young adults, who use it to build and maintain their network of social relationships; about half of all text messages can be categorized as phatic, or social, communication.
- Social media are designed to help people build and maintain relationships; however, frequent use of social media can have a negative impact on subjective well-being. This is in contrast to face-to-face interactions, which generally increase subjective well-being.
- Texting while driving leads to a 400% increase in time spent looking away from the road and a twenty-three-fold increase in risk of accident; texting while walking is dangerous as well, and although texting pedestrians go through the motions of exercising caution, they're still more prone to accidents than nontexters.
- Despite fears that texting leads to declines in spelling and reading abilities, research shows the opposite to be true. The use of homophones and alternate spellings in text messages requires a high degree of phonological awareness; furthermore, habitual texters are sensitive to the situations in which it's appropriate to use textisms.

If you're like most college students, you probably send and receive around forty text messages every day (Morrill et al., 2013). Texting is much cheaper than talking on a cell phone, and oftentimes more convenient. (Just try talking on your cell phone instead of texting in class and see if your professor doesn't notice!) However, texting does have its drawbacks. Unlike making a phone call, which enables you to talk directly with someone a great distance away, texting requires you to enter written language into a machine for transmission. This is generally done by pressing keys to input letters, numbers, and other symbols.

Keyboards provide a separate key for each letter or number, although auxiliary keys like "Shift" and "Control" give the keys alternative functions. More than three dozen keys are needed to represent the letters, numbers and punctuation marks used in English, and so keyboards generally require the use of both hands. You're no doubt acquainted with the QWERTY keyboard used with computers, but some smartphones provide an alphabetically arranged keyboard.

Keypads, on the other hand, are designed for use by a single hand, or even a single finger, and they do this by placing multiple letters on the same key. You're certainly familiar with two different keypad systems, although you may never have noticed before that they're different. The calculator keypad arranges the numbers in descending order, while the telephone keypad arranges them in ascending order. (Go ahead and compare your cell phone keypad to a calculator keypad!)

The point of this discussion is that there's no natural or optimal way to enter text into a machine. Rather, the various keyboard and keypad arrangements we have today are due as much to historical accident as they are to practical considerations.

A Brief History of Texting

The history of texting goes back more than a century and a half with the invention of the typewriter. In the middle of the nineteenth century, various typing machines—each with a different keyboard layout—were developed. However, it wasn't until some decades later that the typewriter and its QWERTY keyboard were standardized (Bi et al., 2012).

According to urban legend, the QWERTY keyboard was designed to slow down the typist to avoid jamming levers, but in fact this is only half true. Letters that commonly co-occur were placed on opposite sides of the keyboard to keep levers from jamming, and as a result skilled typists could enter text much faster that they could on an alphabetic keyboard (Higginbotham et al., 2012).

Even though lever-jamming isn't a problem for computers, the QWERTY keyboard still has its advantages. Skilled typists can type faster when alternating hand and finger movements are required for entering letter sequences (Kozlik et al., 2013). They also report that these sequences are easier to type than sequences using the same hand or finger.

You may have learned touch typing in school, meaning that you can type without looking at the keyboard. If so, you've developed a set of complex and automatic motor skills that enable you to enter text quickly and with relatively few errors. Now that computers have pervaded virtually every aspect of modern life, keyboarding skills are essential.

Your parents and teachers may type faster and more accurately than you because they've had decades more experience on the keyboard. However, you're no doubt faster and more accurate than your elders when it comes to entering text on a keypad (Beeson et al., 2013). Because text messaging is still a new technology, the older generation has no practice advantage, and furthermore they're less likely to send text messages compared with adolescents and young adults. As a result, many young people can touch type on a keypad, oftentimes holding the cell phone in one hand and composing messages with their thumb.

As a memory aid, letters were originally associated with the numbers on the telephone dial, which predated the telephone keypad. The letters provide the opportunity to associate difficult to remember number sequences with easy to remember words, a practice that's still widely used in advertising today. The arrangement of letters on the keypad isn't optimized for inputting text, and cell phone manufacturers have developed a variety of keypad layouts that are more efficient. However, since so many people are already familiar with the existing systems, they're reluctant to spend the time and effort to learn more efficient methods for inputting text (Zhai & Kristensson, 2012). Thus, the traditional telephone keypad, like the QWERTY keyboard, is a historic relic that has become deeply ingrained in our society.

We can think of keypad design as "legacy" technology that isn't well fitted to modern needs. Originally, text was input to a cell phone via **multitap**, which is *a keypad entry method that cycles through the letter values of the key* (Sawyer & Hancock, 2012). Let's say you want to send the simple message *thanks*. In the multitap method, you need to press the 8 key one time for T, the 4 key two times for H, the 2 key one time for A, the 6 key two times for N, the 5 key two times for K, and the 7 key four times for S. That's twelve key presses for six letters.

Since then, newer technology has been developed to make text messaging more efficient. **Predictive text** is *software that uses the probabilities of language to make text entry on a keypad more efficient* (Sawyer & Hancock, 2012). Because no other common word can be spelled with the number sequence 842657, predictive text software will interpret it as *thanks*.

Many times a particular number sequence maps onto multiple potential words. In that case, predictive text software takes into account the preceding few words to make its best guess (Higginbotham et al., 2012). For example, the number sequence 7243 could make at least half a dozen different common English words, such as *page, rage, sage, paid, raid,* and *said*. However, if you've just typed *What she*, then by far the most likely next word is *said*. Predictive software allows people to enter text fluently and within the normal range of language processing speed.

The recent introduction of tablets has presented a new challenge to keyboard design (Trudeau et al., 2013). As mobile computing devices, tablets are generally held in both hands, and a virtual QWERTY keyboard on the touch screen is provided for text entry. However, because of the way the tablet is held, users generally type with their thumbs, as they do on a cell phone. This anatomical limitation has led to the design of the split keyboard when the tablet is held in landscape orientation.

Keyboard and keypad designs are the product of an area of applied research known as **human factors**. This is *an interdisciplinary field that studies the ways that people interact with the machines they use*. We've already seen that human factors generally involves a trade-off between ease of learning and efficiency of use. As a rule, consumers tend to prefer methods that are easy to learn, even if they're inefficient to use.

Texting Today

Some three-quarters of adult cell phone owners also use their mobile devices for sending and receiving text messages (Beeson et al., 2013). This figure no doubt reaches close to 100% for adolescent and young adult cell phone users. About half of all text messages can be categorized as **phatic communication**, that is, *language exchange for the sole purpose of building and maintaining social relationships* (Morrill et al., 2013). The bulk of face-to-face language interaction is also phatic in nature, but mobile technology has also opened the opportunity to maintain social relationships at a distance.

Texting is an inexpensive and private form of communication that is especially well suited to adolescents and young adults, who are actively building their network of social relationships (Conti-Ramsden, Durkin, & Simkin, 2010). However, it lacks the visual and auditory components of talk-in-interaction that convey emotion and underlying meaning (Tulane & Beckert, 2013). As a result, receivers are prone to misunderstanding intended messages, which they incorrectly perceive as offensive.

Although young adults use texting and instant messaging on social media to enhance relationships, research suggests that too much "face time" on Facebook can make you sad (Kross et al., 2013). In a longitudinal study, researchers tracked Facebook use and **subjective well-being** in young adults five times a day for fourteen days. When people engage in face-to-face social interaction, their *personal rating of happiness and life satisfaction* generally increases afterward. However, after interacting on Facebook, the participants' mood tended to become more negative, regardless of what their level of subject well-being was before the Facebook session.

Texting in the classroom has become a contentious issue in high schools and on college campuses. In one survey, both high school and college students believed they were adept at texting in class without their instructors knowing about it (Tulane & Beckert, 2013). Some respondents even reported that they'd observed their teachers texting during class! However, McDonald (2013) found a negative correlation between in-class texting and final grade in a freshman-level social science course. Although GPA, ACT score, and attendance were the three biggest predictors of final grade, in-class texting behavior also had a significant impact on class performance.

Although frequent texting may satisfy a psychological need for connecting with others, it may also have negative health consequences. Researchers monitored two key indicators of stress, heart rate and respiratory rate, of young people as they sent and received text messages. Both heart rate and breathing rate increased as the participants composed messages, and heart rate shot up every time a text message was received. This finding suggests that frequent texters may be subjecting themselves to unhealthy levels of stress.

Texting on the Go

In the last section, we saw how talking on a cell phone impairs driving to a greater degree than being drunk. It only follows then that texting and driving are a bad mix. Simply put, time spent looking at a cell phone screen is time spent not looking at the road. In simulator studies, drivers who texted increased the time they looked away from the road by 400% compared with nontexters (Park et al., 2013).

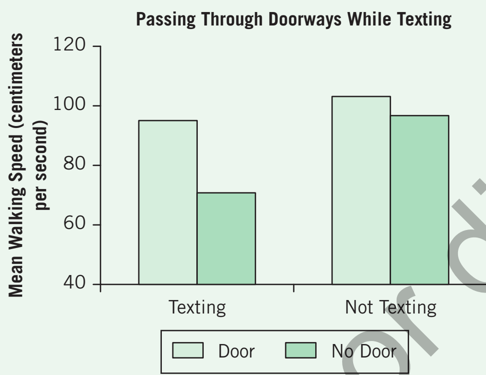
Texting while driving leads to a twenty-three-fold increase in risk of accident, and it's illegal in most states (Sinsky & Beasley, 2013). Still, more than a third of young Americans admit to sending and receiving text messages while they drive on a regular basis (Lopresti-Goodman et al., 2012). Here's a case where people's confidence in their abilities clearly doesn't match up with their actual performance.

Texting while walking is dangerous as well. In one survey, more than a third of respondents admitted to colliding with another person or an object while sending or receiving text messages (Schaburn et al., 2014). Observational studies have found that when people text while walking, they tend to reduce their pace and deviate more frequently from a straight path (Schaburn et al., 2014). Although texters may be slowing down in an attempt to compensate for the distraction of texting, they still put themselves in danger. In simulator studies, texting pedestrians both looked away from traffic more and experienced more hits by motor vehicles than nontexters.

People who text while walking often go through the motions of exercising caution, but in fact they experience more accidents than nontexting pedestrians. In one study, researchers observed how texting and nontexting walkers passed through doorframes of varying width (Lopresti-Goodman et al., 2012). Nontexters were very consistent in their judgments of whether they could walk straight through the doorframes or would have to turn their shoulders to fit through. Texters, on the other hand, rotated their shoulders to pass through doorways that they walked straight through when not texting. They also slowed down more than nontexters.

Figure 13.7 Passing Through Doorways While Texting

Nontexters did not significantly decrease their walking speed as they passed through an experimental doorway. However, the texters did slow down, even though they had plenty of room to pass. They were also more likely to bump into the door frame than were the nontexters.



Source: Based on data from Lopresti-Goodman, Rivera, and Dressel (2012).

Texting impacts safety while walking in other ways, too. Because of the need to hold the mobile device steady in the field of vision, texting pedestrians assume a posture that's not optimal for walking (Schaburn et al., 2014). When we walk, we alternate the swing of our arms to help maintain balance, and we hold our head relatively stable as the rest of our body moves. However, texters hold their upper bodies more rigid than nontexters, which has a negative effect on balance. Texters walk slower in an apparent attempt to compensate for their poor posture, but they're still more prone to mishaps.

CUL8R

There's clearly a generation gap in texting culture. Older adults bemoan the death of the English language at the hands of young texters who have no respect for proper grammar or spelling. However, it's important to keep in mind that languages are living systems that are always in flux. A close inspection of the texting language that adolescents and young adults use shows that it is highly creative and works within established rules (Tulane & Beckert, 2013).

The widespread use of instant messaging technology in the twenty-first century has led to the rise of a new language form that very much parallels the development of creoles and signed language. We've seen in other chapters how children who don't share a common tongue will collaboratively build a language when they're brought together as a social group. The generation that grew up with instant messaging capability have likewise developed an economical and expressive language style based on written English that's well suited to the limitations of the text message format (Anjaneyulu, 2013).

Although text messages are composed in written language, they don't strictly adhere to the traditional conventions of formal writing (Conti-Ramsden et al., 2010). The rules of grammar and spelling are relaxed, and a certain degree of typing errors is tolerated. However, texters are expected to follow the rules of texting etiquette. For example, texters are allowed a certain amount of creative freedom with abbreviations, but these must be transparent to the receiver. *An innovative use of language in text messaging* is known as a **textism**.

Despite anecdotes of textisms creeping into standard English writing as well as concerns about a general decline in the quality of schoolwork, there's little hard evidence to support these claims (Plester, Wood, & Joshi, 2009). In fact, research shows that skill in the use of textisms in instant messaging is positively correlated reading and spelling abilities in the classroom. This is because the use of textisms, with their widespread use of homophones and alternate spellings, requires a high degree of phonological awareness, a necessary precursor to developing strong reading and spelling skills.

An important part of being a fluent speaker (and writer) of a language is knowing how to adjust the style and level of formality to fit the demands of the social context and the expectations of the audience. Using appropriate texting language is simply part of the skill set any communicator in the twenty-first century needs have.

In Sum

Since the invention of the typewriter in the mid-1800s, various schemes have been devised for entering text into machines. Many keyboard and keypad designs are informed by human factors research, which studies the ways that people interact with machines. Taking into consideration the statistical properties of language enables designers to build text-input systems that are efficient both for the machine and for the human operator. Texting on cell phones has become an important means of communication for young people, who use the technology for building and maintaining social relationships. While face-to-face interactions generally lead to an increase in subjective well-being, interactions via social media are more likely to lead to a decrease in happiness and life satisfaction. Texting while driving is exceedingly risky and frequently has lethal consequences; however, texting while walking is also fraught with danger. A new texting culture has arisen that makes creative use of written language. In spite of fears to the contrary, young people who use frequent textisms also have a strong command of spelling and grammar conventions in formal writing.

Review Questions

1. What is human factors research? Discuss some ways that human factors research has informed the design of keyboards and keypads for entering text into machines.
2. What is phatic communication, and what is its relationship to subjective well-being? How do face-to-face and social media interactions differ in terms of subjective well-being?
3. Discuss the dangers of texting while driving and while walking. Based on the results of the doorway experiment (Lopresti-Goodman et al., 2012), do you think it's possible for people to be careful enough to walk and text at the same time? Explain your reasoning.

Thought Questions

1. If you're a regular user of social media, you might have some intuitions about why social media use can lead to a decrease in subjective well-being. What do your intuitions tell you? Can you formulate these intuitions as hypotheses that you could test with experiments?
2. If you know someone who habitually texts while driving, what can you say or do to convince him or her of the dangers? Why do you think people engage in such risky behaviors as texting while driving or walking?
3. You hear an older person complain that text messaging is destroying the English language. How do you respond?

Google It! The Dvorak Keyboard

If you never learned how to touch type on the QWERTY keyboard, you may instead be interested in learning the **Dvorak keyboard**, which many claim allows for faster typing. You can find descriptions and tutorials online. Most computers have Dvorak as an optional keyboard setting.

SECTION 13.3: NATURAL LANGUAGE PROCESSING

- The term *natural language processing* refers to software that responds to, transforms, or produces speech or text; its three main tasks involve (1) converting speech to text, (2) transforming text, and (3) converting text to speech.

- Natural language processing programs use statistical learning mechanisms to extract patterns from a large corpus of text or speech, building up a grammar of the language in much the same way that human infants do.
- Automatic speech recognition refers to the process of converting spoken language input into a text format; common speech-to-text programs include dictation software and programs that output closed captioning for TV.
- A speech-generating device is a computer that converts text into spoken language output, in the form of either prerecorded messages or synthesized speech; assistive reading software is an example of text-to-speech software that is widely used.
- Natural language processing devices that integrate all three tasks (speech-to-text, text-to-text, and text-to-speech) allow for complex human-machine interactions; the personal assistant Siri on Apple's iPhone and Google Translate are examples of interactive natural language processors that are commercially available.
- Many natural language processing devices provide augmentative and alternative communication methods for users who have lost the ability to speak due to disease or injury; while synthetic voices sound unnatural, sometimes individuals can preserve samples of their speech through voice banking.

In the last two sections, we explored how technology has enabled us to communicate *through* machines. In other words, we use machines to convey human language to others who are distant from us in space or time. In this section, we'll consider ways that technology now lets us use human language to communicate directly *with* our machines. **Natural language processing** refers to *software that responds to, transforms, or produces speech or text* (Higginbotham et al., 2012). While talking computers and robots have long been a staple of science fiction, we've now reached an era when talking with our machines, from personal assistants on our smart phones to GPS devices in our cars, has become commonplace.

When HAL Met Siri . . .

HAL 9000 was a computer with natural language processing abilities in the fictional *2001: A Space Odyssey* film and novel series by Arthur C. Clarke. We've long passed the year 2001, when HAL was already operational, and no natural language processing device currently on the market even comes close to HAL's abilities; but in the last few years we've seen significant progress.

Early attempts at developing natural language processing software took the approach of directly encoding the rules of the language into the program. These early programs generally didn't perform very well, or else they only worked within a narrow scope of contexts. In part, the problem was that the structure of language is just too complex to write out as an explicit computer program. Furthermore, all language rules have exceptions, and these need to be encoded as well.

Modern approaches to natural language processing take a statistical approach (Higginbotham et al., 2012). Instead of writing explicit rules, programmers first create a **corpus**, which is *a large selection of text or speech that is used to train natural language processing software*. Oftentimes a training corpus will contain more than ten million words. The software works through the corpus, tracking co-occurrence frequencies and detecting patterns. In other words, natural language processors use statistical learning mechanisms to build up a grammar, just like human infants do. This statistical approach is what accounts for the great strides researchers have recently made in producing practical natural language processing devices for the market.

We can divide natural language processing into three broad tasks: (1) converting speech to text, (2) transforming text, and (3) converting text to speech. These three tasks are combined in various ways to provide useful devices that humans can interact with using natural human language.

Listen to Me

Automatic speech recognition refers to *the machine process of converting spoken language into text* (Zhang, Sun, & Luo, 2014). As symbol manipulators, computers are well suited to processing text, which is composed of a discrete set of alphanumeric characters. It's no surprise that text editors and word processors were among the first computer programs developed, available even in the days of huge mainframe computers before PCs were invented. Recognizing speech, on the other hand, is a far more challenging task for a computer.

Although we perceive the continuous speech stream as being composed of a series of separate phonemes, we learned in Chapter 3 that these speech sounds blend into and overlap one another, and they also take different forms depending on context. Human infants use categorical perception mechanisms to extract the phonemes from the speech stream. Automatic speech recognition software is trained by exposing it to vast amounts of speech along with its corresponding text transcription. As the computer chugs through the data, it learns to associate various speech patterns with particular phonemes (Higginbotham et al., 2012). Trained software can then take speech input and produce an output either in phonetic symbols or in standard written language format.

Speech-to-text programs have a number of useful applications. Depending on the make and model of your laptop or desktop computer, you may already have dictation software installed. If so, you should give it a try. You'll probably find that touch typing is still faster and more accurate. However, you can imagine how convenient this kind of software is for people who lack full use of their hands.

Another application of speech-to-text software is in the production of closed captioning for television programs (Razik et al., 2011). Closed captioning is commonly displayed on TV screens in noisy public locations, such as airports and fitness centers, where the spoken content can't be heard. It also enables deaf and hard-of-hearing people to understand the programs they're watching. In some cases, captioning can be a fully automated process, but usually it requires the assistance of a human moderator. This in part accounts for the time lag between when the lines are spoken and when the captions appear on screen. Despite human moderation, real-time closed captioning can sometimes be difficult to comprehend.

Early automatic speech recognition programs had to be trained on the voice of each speaker that used it (Young & Mihailidis, 2010). This usually involves having the user read a specific script into the computer, which allows it to learn the unique vocal characteristics of the speaker. Dictation software trained in this way can achieve a fairly high level of accuracy. However, other programs, such as customer service applications, need to be able to respond to a wide variety of voices. These programs work well so long as the speech input is limited. If you've ever checked a bank or credit card balance over the telephone, you've had the experience of interacting with automatic speech recognition software.

"Press 1 for . . ."

A **speech-generating device** is *a computer program that produces spoken language output* (Drager, Reichle, & Pinkoski, 2010). Some speech-generating devices play back prerecorded messages spoken by a real human. Those annoying "Press 1 for . . ." telephone menus are of this type. Other speech-generating devices output **synthesized speech**, which is *computer-generated sound output that mimics the fluctuations in fundamental frequencies and formants of human speech*. Synthesized speech allows for more flexibility in how the system responds, since it isn't limited to a set of prerecorded messages. On the other hand, synthesized speech lacks the acoustic detail of real human speech, especially the overtones that provide emotional information, and so it can be difficult to understand.

Synthesized speech is an important component in text-to-speech software. Many computers now have text-to-speech software installed, often with a selection of voices to choose from. If you haven't already, you should try out the text-to-speech program on your computer to hear what it sounds like. You'll probably find the voices monotonous and difficult to pay attention to if they're reading long passages. Nevertheless, synthesized speech is a great boon to those with visual impairments since it enables them interact with their computers, read email, and surf the Internet—activities we take for granted in the twenty-First century but until recently were limited to the visual modality.

People with reading disabilities also benefit from text-to-speech software. **Assistive reading software** is *a computer program that extracts text from a scanned document and then converts the text to synthesized speech*. Studies have found that assistive reading software can help children with learning difficulties improve their reading skill (Chiang & Liu, 2011). Having the computer read aloud as the child reads silently helps the learner develop decoding skills (Chapter 11). Furthermore, children with attention-deficit disorder who use assistive reading software pay more attention to what they read, are less distracted, and read faster with better comprehension. Rather than removing the incentive to read, assistive reading software helps struggling readers develop stronger literacy skills and encourages them to read more.

Q & A

A **question-answering system** is *a computer program that takes natural language inquiries as input, searches databases or the Internet for information, and then provides answers in the form of natural language* (Furbach, Glöckner, & Pelzer, 2010). When search engines were first developed, you could only search the Internet by using keywords. A simple question like *Who was the first president of the United States?* would stump the search engine, because it would not only search for *first president* and *United States* but also *who*, *was*, *the*, and *of*. Later search engines such as Ask Jeeves (now Ask.com) could handle questions in natural language, mainly by filtering out the function words and doing searches on the content words that remained. Recently marketed personal assistants like Apple's Siri on the iPhone are true question-answering systems in that they take natural spoken language as input and produce synthesized speech as output.

Machine translation is another example of a question-answering system (Green, 2011). These programs take text in one language and convert it into text in another language. As we learned in Chapter 9, you can't just translate word-for-word from one language to another. Each language has a unique vocabulary and set of syntactic structures. Again, taking a statistical learning approach is what has driven rapid improvements in machine translation over the last few years.

Let's suppose we want to build an English-Thai translator. First, we need to find a large number of parallel texts in English and Thai, that is, documents in one language that have already been translated (by humans) into the other language. As the computer works through these texts, consisting of many millions of words, it looks for phrase-by-phrase correspondences between the two languages. It then uses these correspondences to translate from one language to the other.

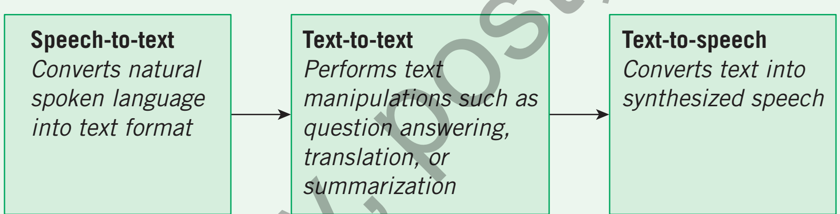
Notice that the machine translator knows nothing about the vocabulary or grammar of either language. Rather, it only knows typical phrase-by-phrase correspondences between languages. This means that a separate translator needs to be trained for each language pair, for example, English-Thai, English-Chinese, Thai-Chinese, and so on. Given that there are some 3,000 languages spoken in the world, we'd need to train some nine million different translation programs to have a true universal translator. While many of the common language pairs now have machine translators on the market, I wouldn't expect to see a Choctaw-Cherokee translator available anytime soon.

Commercially available machine translation software such as Google Translate does a fairly good job with conventional prose such as business letters and instruction manuals. However, translating technical or creative writing still requires a human translator, although machine translation can provide a first draft for the human translator to clean up.

The ultimate goal is to produce a voice-to-voice machine translation system, and rapid progress is being made in that direction (Hyman, 2014). The basic plan is straightforward. First, an automatic speech recognition system takes input in the form of the spoken source language and converts it to text. Next, a machine translation system converts the text in the source language into text in the target language. Finally, a speech synthesis system produces output in the form of the spoken target language. notice that the speech recognition and synthesis programs need to be able to take inputs and produce outputs in both languages. The main hurdle at this point is getting all of these systems to work together seamlessly and in real time.

Figure 13.8 Natural Language Processing Systems

The three classes of natural language processing tasks can be implemented individually or in combination to create all sorts of useful devices.



Giving Voice to the Silenced

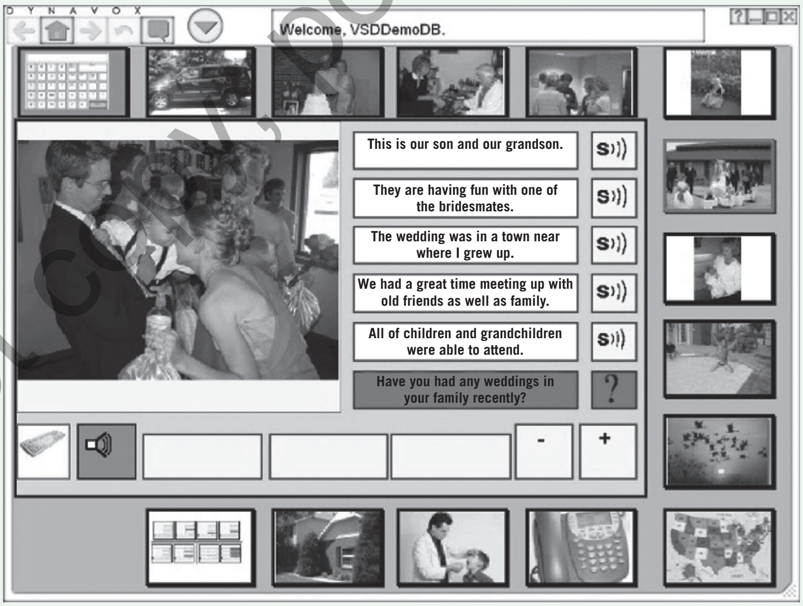
Language is what makes us human. Losing the ability to communicate with others due to brain disease or injury is perhaps the worst possible fate, as it condemns the person to an existence of social isolation. Over 100,000 Americans acquire aphasia each year, mainly due to stroke (Fried-Oken, Beukelman, & Hux, 2012). Although many regain their language faculties, many others don't recover and must rely on technological aids to interact with others. **Augmentative and alternative communication** is *an umbrella term for strategies and devices that help those with language impairments to engage in social interactions*. While the term includes low-tech methods such as note cards and gesturing, it has increasingly come to refer to high-tech devices that supplement or replace a person's ability to speak.

Tablets and smartphones provide convenient platforms for augmentative and alternative communication software. Aphasic patients with intact reading abilities can navigate through the various functions of their device by following written labels. However, aphasia oftentimes disrupts reading as well as speech. In this case, an effective approach is to provide **visual scene displays**, which are *contextually rich and personally relevant digital images that convey content on mobile devices* (Fried-Oken et al., 2012). These naturalistic scenes immediately capture visual attention and are easier to process than are line drawings or abstract symbols (Light & McNaughton, 2012). Each visual scene display has a number of prerecorded messages associated with it.

Stroke can also lead to loss of manual dexterity, in which case pressing buttons on a tablet is impractical. When some speaking ability still remains, as in the case of dysarthria (Chapter 4), using automatic speech recognition software may be a feasible approach (Mustafa et al., 2014). Typically, speech recognition programs don't perform well with dysarthric speech. However, when the software can be trained so that it adapts to the user's particular speech characteristics, satisfactory results can sometimes be obtained. In other cases, a multimodal approach works better (Higginbotham et al., 2012). For example, the user can type the initial letter before speaking the word. The keyboard input then helps the software identify the spoken word input.

Figure 13.9 Visual Scene Display

Visual scene displays on mobile devices enable individuals with aphasia to socially interact with other people.



Source: Fried-Oken, Beukelman, & Hux (2012).

Speech-generating devices can be used to give a voice to those who've lost the ability to speak (Baxter et al., 2012). However, many aphasic patients are reluctant to use synthetic speech to communicate (Khan et al., 2011). Voice is an important component of one's personal identity, and if the user doesn't like the synthetic voice, he or she may feel embarrassed and thus have less motivation to engage in social interaction. In the case of sudden language loss, synthetic speech is the only option; but attempts are made to match the new voice to the original as much as possible.

Sometimes patients still have an intact voice, but they know that they'll lose it soon due to a degenerative disease or impending surgery. In such a case, patients can prepare by going through *the process of recording voice samples for later use in a speech-generating device*. This is known as **voice banking** (Khan et al., 2011). In most cases, the best the patient can hope for is to record a good number of useful sentences that can later be selected and played back as needed.

When sufficient recordings of a person's speech exist, it's possible to create a synthetic voice that sounds quite similar to the original. This was the case with movie critic Roger Ebert, who lost the ability to speak after cancer surgery on his jaw. Because he was a noted TV personality, there were hundreds of hours of recorded speech providing raw material for a speech synthesizer. You can learn more in his TED Talk, available at ted.com.

In Sum

Natural language processing involves any software that responds to, transforms, or produces speech or text. It involves three classes of processes: (1) converting speech to text, (2) transforming text, and (3) converting text to speech. Natural language software uses statistical learning mechanisms to extract patterns from a large corpus of speech or text in much the same way that human infants learn their language. Automatic speech recognition software converts spoken input into text format, and useful applications include programs for taking dictation and for closed captioning TV shows. A speech-generating device converts text into spoken language, either as prerecorded messages or else as synthesized speech. Assistive reading software makes use of synthesized speech. Question-answering systems and machine translation devices incorporate all three natural language processing tasks to create a rich experience of human-machine interaction. Natural language processing in the form of augmentative and alternative communication has given individuals with language loss the opportunity to interact socially in ways they could not have before.

Review Questions

1. What is natural language processing? What approach did researchers first take, and why did it fail? Which approach has proven more successful?

2. Describe the three broad categories of natural language processing. Give some examples of practical devices that employ each process. Also give some examples of devices that combine these tasks.

3. Describe some of the natural language processing devices that have been developed for augmentative and alternative communication in recent years.

Thought Questions

1. Cross-cultural researchers who want to administer the same survey in multiple languages often use a technique known as back-translation to ensure that no meaning has been lost or changed. Let's say we have an English-language survey that we want to administer in Chinese. First, we translate the survey into Chinese, and then we translate it back into English. After that, we compare the original English document against the back-translated document to see if there are any discrepancies. Use the back-translation method to test Google Translate. Try different types of documents to see if some topics or genres work better than others.

2. Current machine translation programs are developed for specific pairs of languages, for example English and Chinese. This means that millions of translation programs will need to be developed, one for each possible language pair. One way to reduce that number is to use a pivot language such as English. Then all we need to do is develop a machine translation program between English and each other language. Thus, to translate from Choctaw to Cherokee, we would first do Choctaw to English and then do English to Cherokee. Based on your experience with back-translation in the previous question, what problems would you expect to encounter? How can you expand on the back-translation method to test the practicality of using a pivot language?

Google It! Remaking a Voice

Search **Roger Ebert TED Talk** to learn more about his heroic struggle to remake his voice after losing his lower jaw to cancer. You'll get to compare his synthetic voice against his original, and you can judge for yourself how good it is.

SECTION 13.4: ARTIFICIAL INTELLIGENCE

- The AI debate centers around the question of whether we can build machines that are intelligent in the same way that humans are; the

Turing test proposes that a machine should be considered intelligent if its ability to hold a conversation is indistinguishable from that of a human being.

- ELIZA was an early program that appeared to have the potential to pass the Turing test; people interacting with the program often treated it as if it were human, in what is known as the ELIZA effect. Programs like ELIZA that can converse within restricted topics are called chatbots.
- The Loebner Prize is an annual award given to the program that comes closest to passing the Turing test; researchers debate which kinds of questions would best distinguish humans from computers, but these often lead judges to misidentify humans as computers, in what is known as the confederate effect.
- Good Old Fashioned Artificial Intelligence (GOFAI) takes the position that intelligence is nothing more than complex symbol manipulation. The Chinese Room argument is intended to show that the GOFAI approach is misguided; the symbol grounding problem asserts that symbols only take on meaning outside of the system in which they are manipulated.
- The embodied cognition approach maintains that machines can only become truly intelligent if they have bodies that can sense and interact with the environment; this has led to research in the development of embodied conversational agents. The view of language as social interaction is important to this approach.
- CAPTCHAs are an example of computers using the Turing test to distinguish humans from bots. Humans can rapidly decipher these distorted letter sequences, but computers cannot; however, humans' ability to do this is still poorly understood.

In the 2013 movie *Her*, a lonely man falls in love with his computer's operating system (Ellison, Jonze, & Landay, 2013). Samantha doesn't just respond to voice commands, she engages him in wide-ranging conversation and expresses her emotions. She also displays personality and conscious self-awareness. Except for the fact that she has no physical form, Samantha is in all other aspects indistinguishable from a human being.

Language has been at the center of discussion on artificial intelligence since the field began in the 1950s. Today we're surrounded by "smart" devices that navigate our cars, find information for us, help us schedule appointments, and even vacuum floors. From the start, however, the Holy Grail of artificial intelligence has been to build a machine that's linguistically and socially indistinguishable in its behavior from a human being (Wallis, 2011). In this section, we go beyond the field of natural language to consider the question: Can machines think?

Talk to Me

In the mid-twentieth century, philosophers were already debating whether machines might someday become as intelligent as humans (Hoffmann, 2010). On the one hand, proponents of artificial intelligence argued that any cognitive processes performed by a brain could, in principle, be performed on a computer of sufficient complexity. On the other hand, opponents of artificial intelligence insisted that brains don't process information that same way that computers do, and thus only brains can give rise to truly intelligent behavior. The crux of the issue centers on your definition of intelligence. *The question of whether we can build machines that are truly intelligent* is known as the **AI debate**, and philosophers still dispute this issue.

British mathematician Alan Turing cleverly sidestepped this philosophical quagmire by instead proposing an operational definition for artificial intelligence (Cullen, 2009). Recall from Chapter 2 that scientists often define abstract concepts in terms of concrete measurements. Turing urged computer scientists to do the same. Specifically, he made *the proposal that a machine should be considered intelligent if its ability to hold a conversation is indistinguishable from that of a human being*. This criterion is now known as the **Turing test**.

Turing justified his test by insisting that we should hold computers to the same standard as humans (Schweizer, 2012). After all, philosophers can't agree on a definition for human intelligence either. Nevertheless, we go through life informally evaluating the intelligence of the people we interact with. How do we do that? Think about it: We judge people as smart or stupid based on their behavior, especially the way they talk. Thus, Turing brought linguistic ability to the forefront of the AI debate.

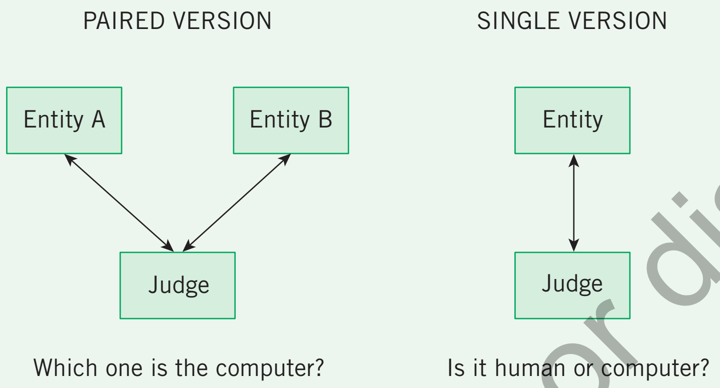
In Turing's day, synthetic speech software hadn't been developed yet, and even today it fools no one for the real thing; so Turing (1950) restricted the format of the conversation to text messaging instead. Imagine you're chatting online with two friends. Now suppose I told you that one of those friends was actually a computer. Do you think you could tell which was the computer and which was the human? Remember, the machine is trying to convince you it's human, so if you directly ask if it's a computer, it will deny it!

The original version of the Turing test works exactly this way. A human Judge communicates with two different Entities. One Entity is a human and is trying to convince you that he or she is human. The other Entity is a computer, and it's also trying to convince you that it's human. After a certain time period has elapsed, we ask the Judge to determine which Entity was the computer. Of course, the Judge has a 50% chance of guessing correctly, but if we run the test on multiple Judges, we can see if group performance is at or above chance. If the group can't reliably distinguish the computer from the human, then the computer passes the Turing test (Pfeiffer et al., 2011).

In an alternative version of the Turing test, each Judge converses with only one Entity, which could be either a person or a machine (Shah & Warwick, 2010). Again, if the Judges perform at chance level, the computer passes the test.

Figure 13.10 The Turing Test

In the paired version, the Judge converses with two Entities and then decides which is the computer. In the single version, the Judge converses with one Entity and then decides if it is a human or a computer.



Passing the Test

Early advances in natural language processing lead to the development of computer programs that appeared capable of passing the Turing test. The most notable of these was ELIZA. Weizenbaum (1966) designed ELIZA to mimic a technique of Rogerian psychotherapy known as reflective listening. Instead of actively engaging in conversation, she waits for the client to write something, and then she responds to that.

ELIZA performs simple grammatical transformations such as changing *I* to *you*. So, if you write *I am feeling a little down today*, she will likely respond with something like *Tell me why you are feeling a little down today*. She also looks for certain keywords, such as terms for family members. So, if you write *My mother really frustrates me*, ELIZA is likely to respond with something like *Tell me about your mother*. She also deftly evades personal questions by insisting that she'd rather talk about you.

If you know in advance that you're texting with a computer, it's fairly easy to see through the simple tricks that Weizenbaum employed in designing ELIZA. However, if you enter into the conversation with the expectation that you're interacting with a human, ELIZA will give you no particular reason to suspect that she's nothing more than software. Certainly she'll make incongruent responses from time to time, but real people chatting on the Internet often spout nonsense as well. And if you're looking for someone to tell your problems to, ELIZA acts more sympathetic than most strangers.

The tendency to misidentify computer-generated text messages as coming from a human is known as the **ELIZA effect** (Shah & Warwick, 2010). The success of ELIZA led to the development of other similar programs. Today, *a computer program that can simulate an intelligent conversation within a restricted topic* is known as a **chatbot**, chatterbot, or conversational agent. In recent years, chatbots have been incorporated into all sorts of natural language programs, including programs that provide customer service online and that serve as virtual characters in video games. However, if you like to chat with strangers online, watch for malicious chatbots posing as real people that try to get you to reveal personal information such as credit card numbers or account passwords.

Eyes on the Prize

So far, no chatbot has been designed that convincingly mimics intelligent conversation over a wide range of topics, as most humans are capable of doing. Since 1991, however, computer scientists have been competing for the **Loebner Prize**, which is *an annual award presented to the computer program that comes closest to passing the Turing test* (Shah & Warwick, 2010).

This contest has spurred a debate about the best approach to take as the Judge in a Turing test. Some Judges take a "power" approach in which they assume the Entity is a computer until they are completely convinced that it's a human. In a study of the transcripts from the 2008 Loebner Prize competition, Shah and Warwick (2010) found that this approach wasn't the most effective. Rather, it tended to lead to what they called the **confederate effect**, which is *the tendency to misidentify human-generated text messages as coming from a computer*. (A confederate is an accomplice of the researcher in an experiment.) It seems that prior expectations strongly affect our perceptions in the Turing test, as they do in the real world.

Another approach is to ask questions that probe emotions or intuitions (French, 2012). For example, if you ask a group of native English speakers whether *Flugbots* is a good name for a new breakfast cereal, nearly all will say no because it just sounds ugly. This kind of question is designed to trip up a computer in a Turing test because it's not likely to have access to this sort of information. This type of question is an acknowledged weakness of the Turing test as a measure of artificial intelligence, since the machine fails not through lack of intelligence but rather lack of human experience (Cullen, 2009).

Back in the Chinese Room

In Chapter 5, we learned about the Chinese Room argument and the symbol grounding problem in the context of understanding how words get meaning. These issues were first presented as potential problems with the Turing test.

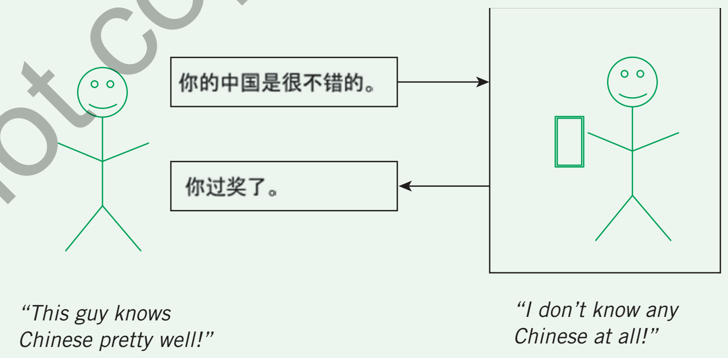
The original version of the Chinese Room argument, Searle (1980), asked us to imagine the following scenario. You're inside a room that has a narrow slit through which people outside the room pass slips of paper with Chinese writing on them. You don't know Chinese, but you have a book that allows you to match each input with the proper response. To people standing outside the room, you appear to know Chinese, but of course you don't. In other words, you passed the Chinese Turing test, even though all you did was manipulate symbols that were meaningless to you. ELIZA and similar programs are all Chinese rooms, according to Searle.

The Chinese Room argument wasn't just intended to challenge the validity of the Turing test. It also attacked *the assumption that many aspects of intelligence can be achieved through the complex manipulation of symbols*. This approach is commonly known as **Good Old Fashioned Artificial Intelligence**, or GOFAI for short (Rodríguez, Hermosillo, & Lara, 2012). According to Searle, the Chinese Room argument demonstrates that mere symbol manipulation can only lead to the appearance of intelligence but not to true understanding. GOFAI supporters conceded that Searle-in-the-room didn't understand Chinese, but some maintained that the room as a whole *did* understand Chinese, an argument that few philosophers or scientists accept nowadays (Ford, 2011).

The Chinese Room argument leads us to the symbol grounding problem, namely the question of how symbols get their meaning (Rodríguez et al., 2012). According to the GOFAI view, symbols acquire meaning through their relationships with other symbols. For example, every entry in a dictionary is defined in terms of other entries in the dictionary, so the whole enterprise seems to be circular, and yet somehow it works. However, English dictionaries are useful to us because we already know many of the words in them, and you certainly can't learn Chinese by memorizing a Chinese dictionary.

Figure 13.11 Setup for the Chinese Room Argument

The Chinese speaker outside the room thinks that the person inside the room must know Chinese. In fact, the person in the room knows no Chinese but only follows instructions in a manual.



As White (2011a, 2011b) points out, symbol systems have both internal and external properties. Internally, symbols have relationships with other symbols inside the system. A dictionary is a symbol system with only internal properties. However, meaningful symbol systems also have external properties. That is, symbols refer to entities or events outside of the symbol system, and this is where their meaning comes from. A dictionary only makes sense if we've already established these external connections between words and their referents.

We still haven't solved the symbol grounding problem. But one approach that's gained currency in recent years is embodied cognition, which we learned about in Chapter 3. In this view, we derive meaning through our interaction with the world. For example, I know what the word *cold* means because I've experienced cold. Since a computer in a box has no experience of the world, it can't understand the symbols it manipulates.

Full Body Workout

The embodied cognition approach has led some researchers in recent years to propose that a truly intelligent machine must have a body with a full range of senses and with freedom to explore and act on its environment (Schweizer, 2012). In other words, it would need to be a robot, not disembodied software. This robot wouldn't be preprogrammed with knowledge of the world but rather it would be outfitted with the ability to learn through experience, just as a human child does. Even then, it's not clear that an intelligent robot's experiences of the world would be the same as those of a human.

Artificial intelligence researchers are also beginning to pay more attention to the fact that language is about much more than just exchanging information (Wallis, 2011). A recurrent theme in this textbook has been the observation that the purpose of most talk-in-interaction is to build and maintain social relationships. Furthermore, meaning in discourse is conveyed not just through words but also—and perhaps even more importantly—through vocal inflections and facial expressions.

Taking a sociocultural view of language use, some computer scientists have turned their attention to the concept of an **embodied conversational agent** (Wallis, 2011). This is *a spoken dialog system that also includes a virtual or physical body*. Many people find computer interactions more natural when they can look at a face that provides them with emotional cues.

Embodied conversational agents currently in use have a limited range of topics and are designed for very specific types of interactions, such as helping online customers place orders. They are also frequently perceived as being cold or condescending (Wallis, 2011). This is because their conversational abilities are still highly scripted. They also don't have the capacity to read their human interlocutor's facial expressions to sense when offense has been taken. Instead of assuaging their partner's hurt feelings, they continue with their scripts, only aggravating the irritation. Machines still have a lot to learn about what makes humans tick.

Turning the Tables on the Turing Test

Humans only account for a portion of online activity. The Internet is full of intelligent software casually known as bots that interact with other bots without any sort of supervision from humans. Bots run the stock market, buying and trading faster than any human can, making fortunes for their masters. They sign up for free email accounts from which to send spam. Countries and companies unleash bots on their enemies and competitors to spy on them. Organized crime rings release bots into social media to flood them with advertisements or to go on phishing expeditions for private information.

When websites need to distinguish between human and computer visitors, they turn to the Turing test. One common test for humanness is the CAPTCHA, which stands for Completely Automated Public Turing test to tell Computers and Humans Apart (Hannagan et al., 2012). Since computers use them to detect humans, CAPTCHAs are sometimes referred to as reverse Turing tests.

CAPTCHAs not only reveal the limitations of the current state of artificial intelligence, they also show us how little we still understand about human perception and cognition (Hannagan et al., 2012). It takes relatively little distortion of an alphanumeric sequence to stump a bot, while humans usually have little difficulty deciphering these twisted letters. Still, it's not clear how people accomplish this task.

Hannigan and colleagues used a priming task to test the hypothesis that our extensive experience with reading in difficult situations, such as reading terrible handwriting, enables us to quickly recognize distorted text (Hannagan et al., 2012). Recall from Chapter 2 that priming refers to enhanced recall due to previous exposure. CAPTCHA primes led to only small decreases in priming compared with printed primes, lending support for the hypothesis.

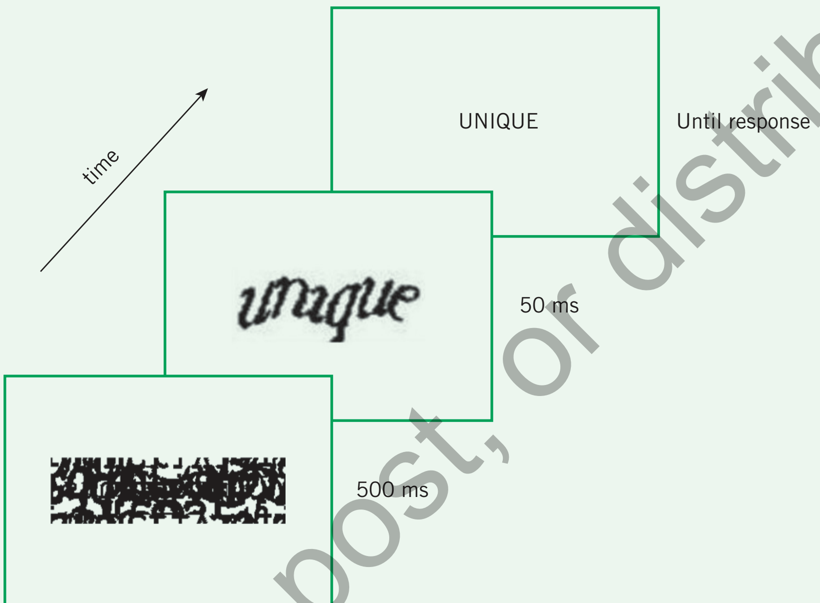
Clearly, computers still have a long way to go before they can match us in the everyday tasks we perform with seeming effortlessness. This observation brings us to the great paradox of artificial intelligence. Tasks that are difficult for humans, such as solving math problems or winning at chess, tend to be easy for computers. Meanwhile, the easy stuff we do without conscious thought still stumps even the most powerful computer.

In Sum

With the invention of the computer in the mid-twentieth century, philosophers and scientists began to take the idea of artificial intelligence seriously. The Turing

Figure 13.12 CAPTCHA Priming Task

Each trial began with a mask for 500 milliseconds, followed by either a CAPTCHA or printed prime for 50 milliseconds. Immediately following that, either a related or unrelated target appeared on screen and remained until the participant responded.

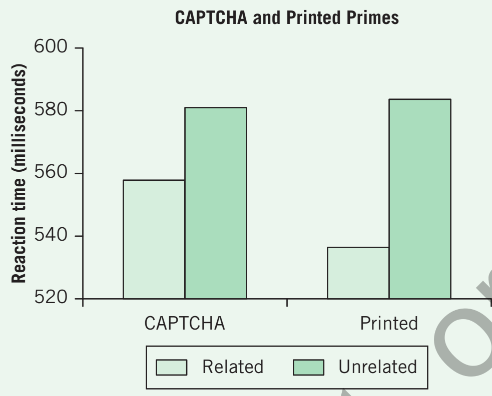


Source: Hannagan et al. (2012).

test considers conversational ability to be the hallmark of human intelligence and the appropriate measure of machine intelligence. Many people willingly respond to ELIZA and other conversational agents as if they were human, and this observation is known as the ELIZA effect. On the other hand, judges in a Turing test often misidentify humans as computers, reflecting the way that expectations guide perception more generally. Good Old Fashioned Artificial Intelligence proposes that intelligence can be achieved through complex symbol manipulation, but the Chinese Room argument refutes that claim. The embodied cognition approach contends that machines can only become truly intelligent like humans if they have bodies that can sense and act on the environment. To protect against attacks by malicious software, many websites require their users to prove themselves as human, often by using a reverse Turing test such as CAPTCHA. Little is known about why humans are so good at deciphering distorted text while computers quickly fail at the task.

Figure 13.13 CAPTCHA and Printed Primes

Reaction times were faster in the Related than the Unrelated conditions for both CAPTCHA and Printed primes. A similar pattern of results was obtained for measures of error rate as well. Since the primes were only shown for 50 milliseconds each, these results support the hypothesis that CAPTCHAs are deciphered rapidly in the early stages of reading.



Source: Hannagan et al. (2012).

Review Questions

1. What is the AI debate? How does the Turing test propose to settle that debate?
2. What is a chatbot? What is the ELIZA effect, and what is it named after? Explain the Loebner Prize and the confederate effect.
3. Explain Good Old Fashioned Artificial Intelligence and its relevance to the Chinese Room argument and the symbol grounding problem.
4. Why are CAPTCHAs considered a "reverse" Turing test?

Thought Questions

1. If you were a Judge in a Turing test, what sorts of questions would you ask? You can even conduct a Turing test yourself by finding a version of ELIZA online. How well does she respond to your questions?

2. Have you ever suspected that someone on a social media site was a chatbot instead of a human? What made you suspicious? Try exploring your intuitions on this.

Google It! Chatbots

Many versions of **ELIZA** are available online. Engage in a conversation with her and try to find patterns that she uses to respond in a seemingly intelligent manner. You can also find plenty of videos on the **Turing test** and the **Loebner Prize** on YouTube. While you're on YouTube, look up some demos of a **chatbot** or an **embodied conversational agent**. Do any of them pass the Turing test?

CONCLUSION

"The report of my death was an exaggeration," wrote Mark Twain in response to the premature publication of his obituary. The same can be said of the English language, which is not only alive but thriving. Pedants and purists may complain that the younger generation is destroying the language with texting slang like ILBL8 and W84M, but nothing could be farther from the truth.

Languages are like living organisms. They're complex systems that are constantly adapting to a changing environment. Only dead languages never change. But then again, nobody speaks a dead language.

All living languages are in constant flux. You don't talk like your grandparents, and they don't talk like theirs. When Shakespeare wrote, "There is a tide in the affairs of men, which taken at the flood, leads on to fortune," he was being poetic but not pretentious. The Bard wrote in the colloquial English of his day, but the language has shape-shifted over the last four centuries.

You, the readers of this book, are the new stewards of the language, the ones who'll remake it to meet the needs of a future we can scarcely imagine. Be bold, be creative. Go ahead and bend the rules—they won't break! After all, the joy of language play is the thing that makes us uniquely human.

CROSS-CULTURAL PERSPECTIVE: Typing and Texting in Other Languages

The QWERTY keyboard was designed with the English language in mind. To prevent typewriter keys from jamming during rapid typing, its inventor placed common letter pairs on opposite sides of the keyboard. While QWERTY is a reasonably efficient layout for typing English, it's not necessarily well suited to other languages. Still, it's become the standard keyboard throughout most of the world.

Shortly after the typewriter was invented in the United States, it was introduced to Europe with only minor tweaks (Bi, Smith, & Zhai, 2012). German-speaking countries swapped the Y and Z keys to make a QWERTZ keyboard. Meanwhile, French-speaking countries made two swaps —A with Q and Z with W—to form the AZERTY keyboard. In both cases, additional keys were added for special letters used in those languages.

Typewriters work well with alphabetic scripts, but not with the logographic systems used in East Asia. There's no straightforward way to accommodate thousands of characters on a two-handed keyboard. As a result, the typewriter never caught on in Japan or China. However, the widespread availability of the personal computer in the 1980s changed all that.

During the twentieth century, both Japan and China had developed romanization schemes for writing their languages in the Roman alphabet, *romaji* and *pinyin*, respectively. There was no intention to replace native scripts, but these schemes do provide a consistent way of representing Japanese or Chinese names in European languages like English.

Figure 13.14 The QWERTZ Keyboard

The QWERTZ keyboard is used in German-speaking countries and other parts of central Europe.

Source: Kozlik, Neumann, and Kunde (2013).

The Japanese learn *romaji* in school, and the Chinese learn *pinyin*; so the letters on the QWERTY keyboard are familiar. All that was needed was a program to convert *romaji* into Japanese script and *pinyin* into Chinese script. In short, the Japanese and Chinese leapfrogged over the typewriter era. In the span of about a decade, they transitioned from cultures where nearly all documents were written by hand to cultures where nearly all documents were composed by word processor.

Typing in Japanese or Chinese is a two-step process. First you type a word or phrase in Roman letters. Then you press a key—typically the spacebar—to tell the computer to convert the letter string into the native script. In the case of homophones, a list of possible equivalents appears, and you press a number to make your selection. Japanese and Chinese word processors also use predictive text software to make good guesses, and so you don't always have to type the entire word or phrase before the software can complete the transcription for you. Texting on a cell phone keypad works similarly. You type in *romaji* or *pinyin*, and software converts it into Japanese or Chinese.

In China, a texting slang has evolved that parallels the textisms of young Americans. Homophonic plays on words are common, swapping Chinese characters with similar pronunciations. As a result, older people often have difficulty understanding the texts of younger folk, just as is the case in the United States.

It's also not uncommon to see letters and numbers mixed in with Chinese-character text messages. For example, MM stands for *mèimei*, written 妹妹 and meaning "sister," while GG stands for *gēge*, written 哥哥 and meaning "brother." These are common terms of address among members of the young, tech-savvy generation in China.

Just as you might text 4U, G2G, or CUL8R, young Chinese texters also use letters and numbers for their sound value. For instance, the number sequence 520 sounds sort of like "I love you" in Chinese. But if you're not ready to commit, it's to time say 886—"goodbye!"

KEY TERMS

- AI debate
- Assistive reading software
- Augmentative and alternative communication
- Automatic speech recognition
- Chatbot
- Confederate effect
- Corpus
- ELIZA effect
- Embodied conversational agent
- Good Old Fashioned Artificial Intelligence
- Human factors
- Inattentional blindness
- Incidental-recognition-memory paradigm
- Loebner prize
- Multitap
- Multitasking
- Natural language processing
- Need-to-listen hypothesis
- P300
- Phatic communication
- Predictive text
- Question-answering system
- Speech-generating device
- Subjective well-being
- Synthesized speech
- Textism
- Turing test
- Visual scene displays
- Voice banking