SPEECH PRODUCTION. The ability to produce speech is the most complex cognitive-motor skill of *homo sapiens.* Wilhelm Wundt was the first psychologist to give a systematic account of this ability and its evolution in *Die Sprache* (Leipzig, 1900). Rudolf Meringer and Karl Mayer, in their *Verlesen und Versprechen* (Stuttgart, 1895), pioneered a major research method, the analysis of speech errors. In 1861, Paul Broca discovered an area in the frontal left hemisphere specifically dedicated to speech production.

The core setting for speech production is conversation, in which participants take turns in contributing to some joint theme. The proximal aim for speakers is that their communicative intentions are recognized from the utterances produced. The construction of such utterances involves several steps, as follows.

Conceptual Preparation

The speaker must select and organize the relevant information to be expressed. If the goal is to give the interlocutor some route direction, the speaker will construct or retrieve a mental map of the appropriate area, decide on a shortest or easiest route, and chunk it in pieces from landmark to landmark, to be delivered in a particular order. The linearization of information to be expressed, to decide on what to say first, what next, and so on, is a core ingredient of conceptual preparation. The selection of information is further guided by the speaker's model of the listener's state of mind. What is the common ground? What information is given? What will be new to the listener? This model of mind is underdeveloped in autistic speech. In order to achieve their communicative goals, speakers have access to a range of rhetorical devices—ways of organizing information effectively—such as asserting, requesting, threatening, and apologizing. The ultimate message constructed by the speaker consists of lexical concepts, concepts for which there exist words in the target language. Languages differ in their repository of lexical concepts, and conceptual preparation differs accordingly.

Grammatical Encoding

To construct an utterance, the speaker will select words (or lemmas) and arrange them in the correct syntactic order. Lexical concepts are part of the mental lexicon, the mental repository of words. Lexical selection is theoretically modeled as activation spreading from a lexical concept node to the corresponding lemma node in a lexical network. In normal fluent speech, lemmas are selected at a rate of two to three per second. Still, selection errors, such as don't *burn your toes* (instead of *fingers),* are rare. They tend to be semantic (as in *toes* for *fingers)* because activation spreads among related concepts. Semantic selection errors are quite frequent in both Broca's and Wernicke's aphasia. The syntactic properties of lemmas drive their syntactic assemblage; they must behave as noun (count, mass, etc.), as verb (transitive, intransitive, etc.), as adverb, as preposition, or as other syntactic category. The utterance is incrementally built up within the constraints of syntactic feasibility. This construction process requires that the syntactic properties of lemmas remain available during some critical time span. Such availability is not guaranteed in Broca's aphasia, typically resulting in agrammatism.

Morpho-phonological Encoding

Each selected lemma spreads its activation to its morphemes in the network. Many words (such as *us, dog, blue,* and *select)* are monomorphemic; others (such as *bill-board, work-ing, eye-s* and *frank-ly)* are multimorphemic. Each morpheme contains a phonological code. It consists mainly of the morpheme's segments, for instance /s/, /i/, /l/, /ε/, /k/ and /t/ for the morpheme select. When a morpheme is successfully activated, the code becomes available. Still, accessing a morpheme can be problematic. In the tip-of-the-tongue (TOT) state, access is blocked, although the lemma can be available with all of its syntactic properties (such as a noun's gender in gender-marking languages). This is also the major problem in certain types of anomia. Generally, accessing a low-frequency morpheme (such as *marsh)* is slower than accessing a high-frequency one (such as *mouth).*

Phonological encoding involves the building up of a syllable structure for the word in its context. When producing the utterance Peter *will select us,* the phrase *select us* will be syllabified as a whole. This proceeds strictly from left to right, as it were, by concatenating the spelled-out segments of the phonological codes: /s/ and /i/ form the first syllable, then /l/, /ε/, and /k/ form the second syllable, and finally /t/, /Λ/, and /s/ form the last syllable: si-lεk-tΛs. Notice that the last syllable, tʌs, straddles a word boundary. This is often the case in syllabification.

Phonological encoding also involves the construction of larger units. Particularly important is the construction of intonational phrases, sense units with a characteristic pitch contour. When asserting *Peter will select us,* the speaker will provide the syllable lɛk with a (raised) pitch accent and then make a dropping boundary tone on the last syllable tʌs. But when questioning the state of affairs, the speaker will render a falling pitch accent on lɛk and provide tʌs with a high boundary tone. The assignment of pitch is strongly affected by the limbic system and hence expressive of emotion.

Phonetic Encoding

As soon as a syllable (such as lɛk or tʌs) is composed, the corresponding articulatory gesture is prepared. We do most of our speaking with no more than a few hundred different syllables. These overlearned articulatory routines are probably all stored in the premotor cortex and retrieved on the fly as syllables are composed. Still, speakers are also able to compose new syllabic gestures for the occasional new or low-frequency syllable. Syllabic routines have highly expressive free parameters for pitch, duration, and amplitude that are independently set during phonological and phonetic encoding.

Articulation

The articulatory gestures are executed by an intricate articulatory apparatus, consisting of the respiratory system, providing the acoustic energy; the laryngeal system, controlling voicing and loudness; and the vocal tract, whose resonance chambers control the timbre of vowels and whose articulators (tongue, velum, and lips) control the place and manner of sound formation. The control of articulation is a major topic in phonetics. [*See* Phonetics.]

Self-Monitoring

Speakers can attend to their own overt and internal speech. When they detect an error or other trouble that may jeopardize the intended communicative effect, they can self-interrupt and make a repair.

[*See also* Aphasia.]

Bibliography

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). New York: Academic Press.

Clark, H. H. (1996). *Using language.* Cambridge, UK: Cambridge University Press.

Fromkin, V. A. (Ed.). (1973). *Speech errors as linguistic evidence.* The Hague: Mouton.

Kent, R. D., Adams, S. G., & Turner, G. S. (1996). Models of speech production. In N. J. Las (Ed.), *Principles of experimental phonetics* (pp. 3-45). St Louis: Mosby.

Levelt, W.J.M. (1989) *Speaking.* Cambridge, MA: MIT Press.

Levelt, W. J. M. (Ed.). (1993). *Lexical access in speech production.* Cambridge, UK: Blackwell.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,*1-38.

Liberman, A. (1996). *Speech: A special code.* Cambridge, MA: MIT Press.

Willem J. M. Levelt