Because the experiments referred to (and others) employ qualitatively different paradigms covering perceptual processes, cognitive evaluations, or movement control, it is proposed that temporal integration up to 3 seconds is a general principle of the neurocognitive machinery. This integration is automatic and presemantic, that is, the temporal limit is not determined by what is processed, but by intrinsic time constants. Because of the omnipresence of temporal integration, it can be used for a pragmatic definition of the subjective present, which is characterized phenomenally by a feeling of nowness, or one can relate temporal integration to singular single states of being conscious (see CONSCIOUSNESS and CONSCIOUSNESS, NEUROBIOLOGY OF).

Additionally, on a different timing level control of motor performance can be registered. Two categorically distinct speed modes with frequencies of 2 Hz and 5 Hz in the sequential initiation of motor behavior are most prominent and can be assessed in simple finger tapping tasks. Nevertheless, they represent basic temporal movement characteristics. Fast automatic movements in the maximum speed in finger tapping can be performed with interresponse intervals of 150 to 200 ms, representing a frequency of approximately 5 Hz. The speed in a personally chosen finger-tapping task is performed with interresponse intervals around 500 ms, representing a frequency of approximately 2 Hz (Fraisse 1982). These two different frequency modes are also seen in other movement tasks and are associated with distinct sensorimotor control processes, the 2-Hz movement being under voluntary control and allowing the collection of somatosensory information, the maximum speed 5-Hz performance requiring only coarse preattentive control (Kunesch et al. 1989). In sensorimotor synchronization where the frequency of a pacer signal has to be reproduced accurately by finger taps, the notion of the categorical difference of the two frequency modes is complemented. The subjective representation of every single finger tap is possible only when a subject is tapping to interstimulus intervals of above 300 ms (Peters 1989). The single taps cannot be temporally resolved in somatosensory perception with interstimulus intervals below 300 ms. This threshold of approximately 300 ms marks the categorical change in motor performance, dividing the aforementioned two motor control processes into automatized movement and voluntarily controlled behavior.

See also INTERSUBJECTIVITY; MOTOR LEARNING; TEMPORAL REASONING

—*Ernst Pöppel*

References

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch, Ed., *The Psychology of Music.* New York: Academic Press, pp. 149–180.

Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology* 35: 1–36

Gomez, C., E. D. Argandona, R. G. Solier, J. C. Angulo, and M. Vazquez. (1995). Timing and competition in networks representing ambiguous figures. *Brain and Cognition* 29: 103–114.

Harter, M. R., and C. T. White. (1968). Periodicity within reaction time distributions and electromyograms. *Quarterly Journal of Experimental Psychology* 20: 157–166.

Hirsh, I. J., and C. E. Sherrick. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology* 62: 423–243.

AuQ: Correct?

Kunesch, E., F. Binkofski, and H.-J. Freund. (1989). Invariant temporal characteristics of manipulative hand movements. *Experimental Brain Research* 78: 539–546.

Mates, J., U. Müller, T. Radil, and E. Pöppel. (1994). Temporal integration in sensorimotor synchronization. *J. Cogn. Neuroscience* 6: 332–340.

Peters, M. (1989). The relationship between variability of intertap intervals. *Psychological Research* 51: 38–42.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences* 1: 56–61.

Pöppel, E., and N. Logothetis. (1986). Neural oscillations in the brain. Discontinuous initiations of pursuit eye movements indicate a 30-Hz temporal framework for visual information processing. *Naturwissenschaften* 73: 267–268.

Schleidt, M., and J. Kien. (1997). Segmentation in behavior and what it can tell us about brain function. *Human Nature* 8: 77–111.

Schwender, D., C. Madler, S. Klasing, K. Peter, and E. Pöppel. (1994). Anaesthetic control of 40-Hz brain activity and implicit memory. *Consciousness and Cognition* 3: 129–147

Tallal, P., S. Miller, and R. Fitch. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. In P. Tallal, A. Galaburda, R. Llinas, and C. von Euler, Eds., *Temporal Information Processing in the Nervous System. Special Reference to Dyslexia and Dysphasia.* New York: The New York Academy of Sciences, pp. 27–47.

Further Readings

Atmanspacher, H., and E. Ruhnau, Eds. *Time, Temporality, Now.* Berlin: Springer.

AuQ: Date?

Block, R., Ed. (1990). *Cognitive Models of Psychological Time.* Hillsdale, NJ: Erlbaum.

Hazeltine, E., L. Helmuth, and R. Ivry. (1997). Neural mechanisms of timing. *Trends in Cognitive Sciences* 1: 163–169.

Nichelli, P. (1993). The neuropsychology of human temporal information processing. In F. Boiler and J. Grafman, Eds., *Handbook of Neuropsychology* 8. Amsterdam: Elsevier Science, pp. 337–369.

Pastor, M. A., and J. Artieda, Eds. (1996). *Time, Internal Clocks, and Movement.* Amsterdam: Elsevier Science.

Tone

One of the fundamental goals of phonological theory is to determine how, starting with a complete phonetic description of a language, we can establish a small number of discrete categories along the various dimensions in which speech sounds may vary (see PHONOLOGY and PHONETICS). These discrete categories along each phonetic dimension are the values assumed by DISTINCTIVE FEATURES, and in any given language they must be rich enough to distinguish utterances that differ with regard to their word choice or their grammatical information, but no richer; other phonetic detail is excluded from such a phonological description. *Pitch* is the name given to the frequency of pulsation of the vocal cords during speech; it describes a continuous variable, which can range from about 50 Hz, at the bottom of a man's range, to about 400 Hz, at the top of a woman's range. *Tone* refers to the small

number of discrete categories necessary for analyzing pitch in a phonological fashion.

Tone is normally analyzed as a property not of individual segments (or sounds), but of syllables. From a descriptive point of view, the pitch of syllables may be either (roughly) constant throughout, or it may have a more complex, dynamic pitch—typically rising or falling. In the first case, such syllables are phonologically assigned a single "level" tone; in the second, in almost all cases, the syllable is analyzed as being assigned a sequence of more than one level tone, and it is this sequence of distinct tones that gives rise to the "dynamic" or "contour" pitch: a sequence of low tone plus high tone gives rise to a rising pitch, and so on. The number of distinct tonal levels (corresponding to different levels of pitch) varies from language to language, but two, three, and four levels of tone are not at all uncommon, and a few languages with even more levels have been reported in the literature. In the neutral intonational pattern of English, a word with initial accent such as "wedding" has a high tone on the first syllable and a low tone on the second syllable, whereas a monosyllabic word such as "man" has a falling pitch, which is to say, the syllable is associated with a sequence of two tones, high and low, in that order.

The term *tone language* is traditionally used to refer to those languages that use differences in tone in order to distinguish between distinct lexical items (that is, distinct words; Pike 1948; Hyman 1978; McCawley 1978; Goldsmith 1994). Such systems may seem exotic from a European perspective, but tone languages are quite common around the world, including a large number of Asian languages (including notably the well-studied Sino-Tibetan languages; Yip 1995), most African languages (especially the well-studied Niger-Congo languages; Odden 1995), and many Mesoamerican languages, especially the Oto-Manguean languages. It is appropriate to bear in mind that in these languages, tone is just as often used to manifest grammatical categories as it is to mark lexical contrasts. For example, in many Bantu languages of Africa, the verb will consist of six or more syllables, composed of a polysyllabic stem with several prefixes. The first syllable of the stem is the lexical root, and its tone is a lexical (that is, idiosyncratic) property. The tones of the following syllables of the stem, however, will be determined by the tense of the sentence, and they are thus considered to express grammatical information.

Languages are often divided into three categories, from a prosodic point of view: tone languages, pitch-accent languages, and intonational languages, but the boundaries between these categories can be difficult to discern at times (see PROSODY AND INTONATION). Pitch-accent systems share in common with tone languages relatively strict principles determining the tone assigned to each syllable, and thus the pitch at which each syllable will be uttered, but rather than utilize tonal differences as such in the LEXICON, pitch-accent systems characteristically employ a single tonal melody (e.g., low-high-low, constant across the language), and assign it to an entire phonological word. However, such languages employ an assignment algorithm that recognizes that one syllable is prosodically special, the syllable that is called "accented." The accented syllable bears a special mark in the lexicon, and tone mapping will typically make that accented status be phonetically unique in some fashion, by making it the final high of the word, or the first high, or through some other means. The information that is thus transmitted by the pitch pattern is which syllable is accented, rather than what the tonal melody is. The location of the pitch accent will in some languages be predictable by simple rule, in others by complex rules involving the interaction of the morphemes involved, and in yet others be simply stipulated in the lexicon. Tokyo Japanese is a familiar example of a pitch-accent system. However, some systems exist (including other dialects of Japanese) that contain two distinct tonal patterns, and lexical items must be marked as to which tonal pattern is appropriate. Systems of this sort share functional characteristics with tone languages, and blur the boundary between tone languages and pitch-accent languages.

Intonational languages (a category that includes English) generally allow pragmatic factors to influence the tone and the pitch of an utterance to a greater degree than would a tone or pitch-accent language. Most intonational languages (French is an exception) share with pitch-accent systems the characteristic of marking exactly one syllable as accented, and using the location of that accented syllable in the algorithm that maps an intonational pattern to the phonological word. However, it is very difficult—indeed, at this point, it seems impossible—to draw a sharp boundary between those aspects of intonation that can be adequately handled with discrete linguistic categories and those that blend gradually into the noncategorical aspects of general communicative behavior. A consideration of the intonation of ironic, polite, or highly emotional speech quickly encounters this problem.

As these remarks suggest, one of the most important characteristics of tonal patterns in language is their considerable autonomy with respect to the other characteristics of spoken language, which are produced largely by the mouth and nose. This autonomy is manifested in several ways, including: (1) the numerical mismatch, or many-to-many relationship, found between tones and syllables; (2) the retention of a tone despite the synchronic or diachronic loss of a tone (so-called tonal stability); and (3) the existence of morphemes that consist solely of one or more tonal specification (so-called floating tones). These observations are closely linked to the foundations of autosegmental phonology (Goldsmith 1990). In addition to the numerical mismatch between tones and syllables, tone languages often display a temporal mismatch between the tones and the syllables of a given morpheme. The consequence of this is that it is not at all uncommon to find in tone languages that a morpheme is underlyingly (or analytically) composed of both tones and syllables, and yet the tone(s) of the morpheme will be consistently realized on a different syllable, either earlier or later than the syllable(s) pertaining to that morpheme.

See also MORPHOLOGY; PROSODY AND INTONATION, PROCESSING ISSUES; STRESS, LINGUISTIC; TYPOLOGY

—John Goldsmith

References

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology.* Oxford: Blackwell.

AuQ: Location?

Goldsmith, J. (1994). Tone languages. *Encyclopedia of Language and Linguistics.* Pergamon Press.

Goldsmith, J., Ed. (1995). *The Handbook of Phonological Theory.* Oxford: Blackwell.

Hyman, L. (1978). Tone and/or accent. In D. J. Napoli, Ed., *Elements of Tone, Stress, and Intonation.* Washington, DC: Georgetown University Press, pp. 1–20.

McCawley, J. (1978). What is a tone language? In V. A. Fromkin, Ed., *Tone: A Linguistic Survey.* New York: Academic Press, pp. 113–131.

Odden, D. (1995). Tone: African languages. In J. Goldsmith, Ed., *The Handbook of Phonological Theory.* Oxford: Blackwell, pp. 444–475.

Pike, K. (1948). *Tone Languages: A Technique for Determining the Number and Type of Pitch Contrasts in a Language, with Studies in Tonemic Substitution and Fusion.* University of Michigan Publications in Linguistics, no. 4. Ann Arbor: University of Michigan Press.

Yip, M. (1995). Tone in East Asian languages. In J. Goldsmith, Ed., *The Handbook of Phonological Theory.* Oxford: Blackwell, pp. 476–494.

Further Readings

Beckman, M., and J. Pierrehumbert. (1986). Intonation structure in Japanese and English. *Phonology Yearbook* 3: 255–309.

Bolinger, D. (1985). Two views of accent. *Journal of Linguistics* 21: 79–123.

Duanmu, S. (1996). Tone: An overview. *Glot International* 2: 3–10.

Fromkin, V., Ed. (1978). *Tone: A Linguistic Survey.* New York: Academic Press.

Hyman, L., and R. Schuh. (1974). Universals of tone rules: Evidence from West Africa. *Linguistic Inquiry* 5: 81–115.

Pulleyblank, D. (1986). *Tone in Lexical Phonology.* Dordrecht: Reidel.

van der Hulst, H., and N. Smith, Eds. (1988). *Autosegmental Studies on Pitch Accent.* Dordrecht: Foris.

Top-Down Processing in Vision

Perception represents the immediate present, what is happening around us as conveyed by the pattern of light falling on our RETINA. And yet the current pattern of light alone cannot explain the stable, rich experience we have of our surroundings. The problem is that each retinal image could have arisen from any of a vast number of possible 3-D scenes. That we rapidly perceive only one interpretation tells us that we see far more than the immediate information falling on our retina. The highly accurate guesses and inferences that we make rapidly and unconsciously are based on a wealth of knowledge of the world and our expectations for the particular scene we are seeing. The influences of these sources beyond the images on the retina are collectively known as top-down influences.

Both top-down analyses and the complementary bottom-up processes use local cues to assign depth to the regions of an image. They differ in the manner in which they resolve the ambiguity of the local cues. A bottom-up analysis, part of MID-LEVEL VISION and SURFACE PERCEPTION, makes direct links between local geometrical features and depth. For example, whenever one object partially covers another, the visible contours of the more distant object terminate at the outer boundary of the nearer one, forming what are called T-junctions. When a T-junction is encountered in an image, this logic can be reversed: the stem of the T is designated a contour of a more distant, partially hidden object and the top of the T is assigned to the outer boundary of a nearer object.

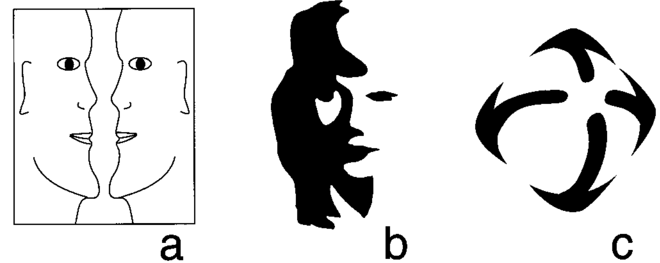

Figure 1.

A top-down process, on the other hand, depends on the content of the image and its analysis by processes of HIGH-LEVEL VISION. Cues operate by suggesting objects—a nose contour might suggest a face, for example—and then stored information about that object's structure can be applied to the assignment of depth in the image. Other features in the image are then examined to verify or reject the postulated object. The cues used for the initial selection of potential objects are not limited to the current images but include preceding images as well as nonvisual sources which affect our expectations for the scene. The sources of object knowledge which are called upon may be built up over both evolutionary or individual time scales.

Our guesses for appropriate internal models are best when we know what to expect in a scene. Upon opening a door to a classroom, for example, we expect to see desks and a black or white board. If these elements are present in the scene, they are rapidly interpreted. Incongruent elements are seen less reliably as Biederman (1981) showed when he reported increased errors in identifying fire hydrants presented in kitchens or sofas floating over city streets than when they were presented in their usual contexts. As Biederman's example demonstrates, top-down analyses work because there is a great deal of semantic redundancy in the content of a scene—noses are expected to be seen along with mouths, cars with roads, classrooms with desks, and sofas with coffee tables; moreover, noses, cars, and sofas have typical shapes so that once a few distinctive features have implied the presence of say, a car, the other expected features of a car can be verified or even just assumed to be present.

Textbook examples of top-down processing typically make use of images with two or more equally likely interpretations which are sometimes referred to as ILLUSIONS. A hint as to which interpretation to see may then trigger one or the other, as in the examples shown here. (a) Two faces, or one vase, or one face behind a vase (Costall 1980); (b) a man playing a saxophone seen in silhouette, or a woman's face in sharp shadow (Shepard 1990); and (c) a sphere in a four-point setting or a white angel (Tse 1998). In these instances,