to omit arguments if they are highly informative and can be recovered from the discourse context. Interestingly, although children have a universal tendency to omit highly informative arguments, there are cross-linguistic differences in the frequency of null arguments in child language that reflect the prevalence of null arguments in the adult input. Studies of children's referential choices in contexts in which they need to identify referents for the benefit of a naïve listener, as for example in a narrative task, have shown that systematic use of discourse-appropriate markers of newness is late, around 6 years of age.

Although corpus studies have shown early sensitivity to discourse-pragmatic constraints in referential choice, experimental studies have provided a more nuanced picture of what children can and cannot do when the addressee's access to a referent is manipulated to exclude it from the common ground. Referential communication games, in which children have to identify a referent for the benefit of an addressee who may or may not have seen the intended referent, demonstrate that children up to the age of 6 cannot systematically use maximally informative referential expressions for referents that are not in the common ground.

Similar results are reported in comprehension studies: Using eye-tracking, Nicholas Epley, Carey Morewedge, and Boaz Keysar observed that children (up to the age of 6) were likely to look first to a referent that matched the interpretation of a referential expression from their own perspective, even when that referent was not visually accessible to the speaker of the utterance. Studies comparing the looking behavior of children and adults suggest that such egocentric interpretations persist for longer in less mature speakers.

Narrative studies, where the main focus is predominantly on children's establishment and maintenance of referential cohesion in a monologic rather than dialogic context, have also provided evidence that the acquisition of local and global markings for referent introduction in narrative contexts is a protracted process that continues well into the primary school years. Children tend to have difficulties using discourse-appropriate expressions to introduce new referents to a naïve listener and use less informative expressions than would typically be required.

While sensitivity to discourse-pragmatic constraints emerges relatively early, mastery of the mapping between discourse status and referential expression is a process that is protracted into the primary school years.

Ludovica Serratrice
University of Manchester

See Also: Alignment (Interactive); Conversational Skills; Gestures in Communicative Development (Overview), Joint Attention and Language Development; Narrative Development; Perspective Taking in Communication; Pointing; Pragmatic Development; Pronouns; Referential Communication; Speech Acts.

Further Readings

Allen, Shanley, Barbara Skarabela, and Mary Hughes. "Using Corpora to Examine Discourse Effects in Syntax." In *Corpora in Language Acquisition Research: Finding Structure in Data,* Heike Behrens, ed. Amsterdam: John Benjamins, 2008.

Clark, Eve. "Pragmatics and Language Acquisition." In *Handbook of Pragmatics,* Laurence Horn and Gregory Ward, eds. Oxford, UK: Blackwell, 2004.

Epley, Nicholas, Carey Morewedge, and Boaz Keysar. "Perspective Taking in Children and Adults." *Journal of Experimental Social Psychology*, v.40 (2004).

Greenfield, Patricia and Joshua Smith. *The Structure of Communication in Early Language Developmen*t. New York: Academic Press, 1976.

Hendriks, Petra, Charlotte Koster, and John Hoeks. "Referential Choice Across the Life Span: Why Children and Elderly Adults Produce Ambiguous Pronouns." *Language and Cognitive Processes* (2013).

Hickmann, Maya. *Children's Discourse. Person, Space and Time Across Languages.* Cambridge, UK: Cambridge University Press, 2003.

Matthews, Danielle, Elena Lieven, Anna Theakston, and Michael Tomasello. "The Effect of Perceptual Availability and Prior Discourse on Young Children's Choice of Referring Expressions." *Applied Psycholinguistics*, v. 27 (2006).

Distributional Knowledge and Language Learning

One of the hotly debated topics in language acquisition is how children learn the structure of the language they are exposed to. To acquire language, children have to learn the linguistic units of their language (phonemes, morphemes, and words) and the productive, yet restricted, ways those units can be combined. Both the parts and the rules differ across languages and are not explicitly taught. The need to simultaneously learn both aspects introduces a serious challenge: How can children discover the regularities without first knowing what the parts are? And how do they extract the correct parts without knowing their defining characteristics? For example, Hindi and English both have aspirated sounds, but aspiration is a phonological contrast only in Hindi. Infants learning Hindi and English have to make different generalizations without knowing in advance what they are looking for (e.g., phonemes) and what the distinguishing properties are (e.g., aspiration).

Over the past 20 years, much research has shown that children can use distributional information–knowledge about the way linguistic elements co-occur—to discover linguistic structure. A seminal study to demonstrate children's ability to use such information was in the domain of speech segmentation. Unlike written language, words in spoken speech are not consistently separated by pauses. One of the cues to an upcoming word boundary is distributional: Transitional probabilities between syllables (the likelihood of one syllable following another) are higher within words as compared to across word boundaries. Using an artificial language—where input can be fully controlled—Saffran and her colleagues showed that 8-month-old infants can use the transitional probabilities of adjacent syllables to segment speech into words. After a two-minute exposure to a novel, unsegmented speech stream, infants posited word boundaries between syllables that had low transitional probabilities. More recent studies have looked at a range of cues infants use to segment speech, including language-specific stress patterns, phonotactic constraints on word-final and word-initial syllables, and articulatory cues.

Distributional information is relevant for learning other aspects of language structure such as its inventory of speech sounds, grammatical categories, and syntactic relations. Phonetic features differ in their distribution when they are contrastive (e.g., aspiration in Hindi) and noncontrastive (e.g., aspiration in English). Infants are sensitive to this difference and make generalizations accordingly: They posit two categories when exposed to a bimodal distribution (most tokens at the two ends of the continuum) and only one when exposed to a unimodal distribution. Infants can learn about word order and head directionality in their language by distinguishing function and content words based on their different frequency distributions (without knowing any of their other properties). Distributional information can also help children form grammatical categories. In the 1980s, M. Maratsos and colleagues argued that grammatical categories are (in part) defined by their distributional properties and that children learn them by grouping together words that co-occur in similar syntactic and morphological environments. More recently, Toben Mintz showed that grouping together words that appear in the same frequent frame (e.g., *you* ___ *it*) leads to the formation of categories that are a close approximation of grammatical categories, like verb or noun; furthermore, adults are able to make use of such analyses to form categories in an artificial language. Similarly, children can use information about the distributional properties of different verb classes (e.g., animate versus inanimate subjects and number of arguments) to assign meaning to novel verbs.

Computational Models

Computational models provide useful ways of studying the role of distributional information in language learning. Such models are used to uncover the distributional correlates of various structural properties using large data sets and to evaluate the reliability of generalizations made on the basis of distributional information. M. Redington, N. Chater, and S. Finch, for example, created a model that uses distributional cues to learn syntactic categories from child-directed speech. Their model uses bigram statistics to group together words that appear in similar distributional contexts, resulting in a hierarchical structure with clusters at various levels (e.g., pronoun cluster in a larger noun cluster). The model's classification is then compared to other computational models and to human judgments to see how reliable and useful distributional cues are in forming syntactic categories. Such models allow us to address multiple questions about the use of distributional information (e.g., Is learning better from child-directed speech? Are distributional cues similarly effective for function and content words?).

Studies of artificial language learning allow us to further probe children's ability to use a variety of distributional cues. Studies ask how sensitivity to different cues changes with age and experience. In word segmentation, for instance, younger infants rely more on statistical cues (transitional probabilities), while older infants assign more weight to stress: a pattern that reflects a move from reliance on universal cues (transitional probabilities are a cue to word boundary across languages) to language-specific ones (stress patterns differ between languages). Other studies focus on the kinds of distributional relations children can (and cannot) attend to. Language has dependencies that hold between nonadjacent words, like the dependency between auxiliaries and inflectional morphemes *(is running)*. Rebecca Gomez's work shows that children are able to learn such relations in an artificial language but only when there are many different intervening elements (like the many different verbs appearing between the auxiliary and the *–ing* morpheme). The combined experimental and computational findings show that many aspects of language structure have distributional correlates and that these distributional patterns are attended to by human learners.

In the past, the main question was whether infants track distributional information at all. Now that infants' ability to do so has been well documented, the focus is on how this information is used in actual language learning. How do infants integrate the different cues they attend to? Experimental studies often present infants with only one distributional cue, but real-life learning involves multiple cues, some of which may point to different generalizations. Behavioral studies and computational simulations are used to ask how infants detect, weigh, and integrate different sources of distributional information. Much of our knowledge about the use of distributional information comes from artificial language learning studies. Yet, it is not clear how well that maps onto real language learning and use. Examining correlations between performance on artificial language learning tasks and language processing is one way to address this question. Another open question is the degree to which the use of distributional information differs in humans and nonhumans. By comparing distributional learning in humans and nonhumans, researchers hope to uncover the unique aspects of human language learning. Additional research asks how inductive biases affect what is learned from distributional information: What makes humans prefer certain generalizations over others, and how is this is related to the way languages have evolved over time?

Conclusion

To conclude, numerous studies have documented the role of distributional information in language learning. The challenge for future research is to spell out the mechanisms that allow children to reach the correct generalizations while integrating and attending to the multiple cues found in real-life learning.

Inbal Arnon
Hebrew University

See Also: Auditory Sequence/Artificial Grammar Learning; Computational Models of Language Development; Frequency Effects in Language Development; Grammatical Categories; Induction in Language Learning; Metrical Segmentation; Multiple Cues in Language Acquisition; Statistical Learning; Word Segmentation.

Further Readings

Gómez, Rebecca L. "Variability and Detection of Invariant Structure." *Psychological Science*, v.13/5 (2002).
Maratsos, Michael. "Are Actions to Verbs as Objects Are to Nouns?" *Linguistics*, v.28/6 (1990).
Mintz, T. H., et al. "The Distributional Structure of Grammatical Categories in Speech to Young Children." *Cognitive Science*, v.26 (2002).
Redington, M., N. Chater, and S. Finch. "Distributional Information: A Powerful Cue for Acquiring Syntactic Categories." *Cognitive Science*, v.22 (1998).
Romberg, A. R. and J. R. Saffran. "Statistical Learning and Language Acquisition." *Wiley Interdisciplinary Reviews: Cognitive Science*, v.1/6 (2010).
Saffran, J. R., et al. "Statistical Learning by 8-Month-Old Infants. *Science*, v.274 (1996).

Domain Specificity

Domain specificity in cognitive science refers to the specificity of the substrates of a particular cognitive domain, such as language. That is, if language is domain specific, at least certain of its underpinnings in the mind or brain are dedicated to language (or aspects of language) and subserve nothing else. In contrast, if language is not domain specific, this means that its substrates underlie not just language but other domains as well. This is usually referred to as domain generality or domain independence. More generally, the question as to whether or in what ways language (or other aspects of cognition) might or might not be domain specific is referred to as the issue of domain specificity. It is difficult to demonstrate