The syntax of mental imagery

Nishant Mulay

Email: nishantmulay@gmail.com

An introspection into one's own mind would plausibly suggest that, fundamentally, two kinds of thoughts arise in the mind – sentences and images. Sentences are formed by virtue of recursive binary set formation (Merge). Lexical items exist, and syntax produces sentences out of them. The topic of how these lexical items are acquired by the human throughout the period of language growth is an interesting question, which has interesting dimensions to it too such as the distinct acquisition of nouns and verbs, but the topic this paper will tackle is the nature of the generation of thought that is images, or the generation of image-thought, if you like.

Imagine this, that there is a homunculus man sitting inside your head, like a lone person sitting in a dark theater (maybe for a midnight show for which he is the sole watcher), enjoying a stream of images projected on the screen before him (and also listening to the corresponding dialogues, monologues, etc. More on this sound part later). The screen that this homunculus man sees is the internal screen of the mind, on which mental imagery is projected, displayed.

To stretch this analogy a bit and develop the main idea this paper wants to talk about, the projector (which only concerns itself with image-generation), is hypothesized to be based in the language module of the brain. That this projector is a device that piggybacks on the generative device that is fundamentally used for language is what this paper mainly hypothesizes.

Jerry Fodor famously proposes a 'Language of Thought', that there exists a language for the internal phenomenon of thought (Mentalese) like there is language for external communication. Noam Chomsky, Otto Jespersen, William Dwight Whitney, Bhartrhari (ancient Indian grammarian), go further and say that there is, essentially, no difference between the internal phenomenon of thought and external phenomenon of, let's say, English. When Bhartrhari says that thought anchors language and language anchors thought, and if he tackles the question of mental imagery, this theory of his will lead him to (if he accepts our assumption at the beginning that mental imagery is a kind of thought) the null hypothesis that mental imagery is anchored in thought, and thought is anchored in language, and so, he will finally state that language anchors mental imagery.

Fodor, when proposing Mentalese, might not have had image-thought in mind, but only the sort of purely logical language of internal thoughts, which comprises things like assertions, inferences, etc. But whoever posits a 'Language of Thought' should also account for the language of mental imagery.

To come back to our analogy of the projector, this projector is what generates mental imagery that our "inner eye", or the eye of this homunculus man, sees. If we make another educated assumption that sentences formed in the mind, the other kind of thought, are fundamentally represented phonetically, i.e., sentences are "heard" internally, then we can say that our projector is in fact attached to a speaker (the electronic kind), and we would now be dealing with a speaker-projector device. Based on this developed analogy, what this paper essentially says is that the projector that projects onto the mental screen is piggybacking on a generative sort of a speaker that potentially generates an infinite set of sentences.

It shouldn't go unnoticed here that what we are also implying is that this speaker is actually the main seat of generative grammar, and the projector is only piggybacking on it, Or, to put it colloquially, hearing takes precedence over seeing, at least when talking about internal structures within the mind.

To consolidate the previous paragraphs into one whole picture, a lone homunculus man is sitting inside our head, sometimes hearing what this generative electronic speaker is speaking, sometimes watching what the piggybacking projector is projecting, and sometimes doing both, watching and hearing a whole movie (comprising sound and image). To further resolve this analogy, when thoughts come to us in sentences, this projector goes dead, and we only hear the speaker. When thoughts come to us in images, the speaker goes mute, and we only see the projection. When both kinds of thoughts simultaneously occur to us, it is like our homunculus man is watching a movie with full sound.

Our posited speaker-projector device gives way to a potentially rich topic for exploration, that of studying the exact nature of this link between sentences and images, between the generative speaker and the piggybacking projector, so to speak. The mechanisms that go into this sort of a transformation of sentences to images is what the topic is. Essentially, this topic asks the fundamental question of what is the nature of the link between the syntax of sentences and the generation of mental imagery. In the following paragraphs, we will try to tease out what we can about this link.

Syntax works on the lexicon. Without a lexicon, syntax does nothing. Without water in the dam, the arrangement of pipes that manage and provide water to the city sit idle. Any kind of syntax needs some atoms (words) for it to work on. After the acquisition of a sizeable lexicon in the human mind, the language-specific generative syntax works on it and potentially produces an infinite set of sentences. Suppose, a sentence is generated in English: "The camel sings, grows wings, and flies away." This sentence, though physically impossible to realize, can still have a complement image, maybe something out of a fantastical cartoon show. This is the type of (comprehensible) sentence that is immediately realizable on the mind-screen. The second sentence, "colorless green ideas sleep furiously" is the type of non-incomprehensible sentence that is not immediately (maybe never) realizable on the mind-screen, at least if you take the dictionary meaning of these words. A third, non-sentence, "colorless furiously ideas green sleep", is incomprehensible and hence provenly never realizable on the mind-screen.

In this third example, English syntax doesn't ever produce such a sentence, and hence there is no representation of it in any form. In the second sentence, English syntax does produce it, as in the structure "Adj + Adj + N + V + Adv" is possible in English, but there is no making head or tail of the sentence in terms of the established dictionary meanings of the words involved. If something like Wittgenstein's builders' community gives different meanings to these words, then the sentence can possibly be comprehensible within this community, but that will just be a case of a dictionary update, a trivial update of the mapping between sound and meaning, and hence pondering on this scenario would be futile.

If we have to tease out a rough sort of a dependency between our posited projector and the generative speaker, then only those sentences that are completely, clearly comprehensible within a language have a clear mental image, and hence can be adequately projected.

Lastly, it would not, given all the above, be absurd now to suggest that the lexical items, along with having linguistic features, also have pictorial features. The exact nature of this link between the pictorial form and the linguistic form of a sentence, and the nature of the pictorial features of a lexical item, are two potential topics of future study.