A new study from MIT cognitive scientists has determined just why legal documents such as contracts or deeds are often so impenetrable. Credit: MIT News

Legal documents, such as contracts or deeds, are notoriously difficult for nonlawyers to understand. A new study from Massachusetts Institute of Technology (MIT) cognitive scientists has determined just why these documents are often so impenetrable.

After analyzing thousands of legal contracts and comparing them to other types of texts, the researchers found that lawyers have a habit of frequently inserting long definitions in the middle of sentences. Linguists have previously demonstrated that this type of structure, known as "center-embedding," makes text much more difficult to understand.

While center-embedding had the most significant effect on comprehension difficulty, the MIT study found that the use of unnecessary jargon also contributes.

"It's not a secret that legal language is very hard to understand. It's borderline incomprehensible a lot of the time," says Edward Gibson, an MIT professor of brain and cognitive sciences and the senior author of the new paper. "In this study, we're documenting in detail what the problem is."

The researchers hope that their findings will lead to greater awareness of this issue and stimulate efforts to make legal documents more accessible to the general public.

"Making legal language more straightforward would help people understand their rights and obligations better, and therefore be less susceptible to being unnecessarily punished or not being able to benefit from their entitled rights," says Eric Martinez, a recent law school graduate and licensed attorney who is now a graduate student in brain and cognitive sciences at MIT.

Martinez is the lead author of the study, which appears in the journal Cognition. Frank Mollica, a former visiting researcher at MIT who is now a lecturer in computational cognitive science at the University of Edinburgh, is also an author of the paper.

Translating legalese

While a student at Harvard Law School, Martinez became interested in how lawyers and judges use language to communicate. He cross-registered for an MIT linguistics class taught by Gibson, and after he finished his law degree, he joined Gibson's lab as a grad student.

In their new study, Martinez, Mollica, and Gibson set out to figure out just why legal documents—such as terms of service agreements, mortgage documents, and other kinds of contracts—are so
hard to understand. First, they compared a large selection of legal contracts (totaling about 3.5 million words) to other types of writing, including movie scripts, newspaper articles, and academic papers.

Using a text analysis tool that can identify patterns in large volumes of text, the researchers identified several features that occur much more commonly in legal documents than in other kinds of writing. As one example, they found that legal documents include many instances of nonstandard capitalization, such as using all caps. This is legally required in some types of legal documents to make certain provisions more conspicuous. Legal writing also features much more use of the passive voice, the researchers found.

When the researchers asked nonlawyers to read either legal documents or documents in which certain features of the text were altered without changing the meaning, they found that the passive voice and nonstandard capitalization did not make the documents more difficult to understand. The biggest culprit, they found, was center-embedding. In this type of construction, a writer introduces the subject of a sentence, then inserts a definition of the subject, and then continues on with the sentence. In their paper, the researchers included this sentence, with a lengthy definition in parentheses, as an example:

"In the event that any payment or benefit by the Company (all such payments and benefits under Section 3(a) hereof, being hereinafter referred to as the "Total Payments"), would be subject to excise tax, then the cash severance payments shall be reduced."

The paper offers this as a more understandable alternative, with the definition separated out:

"In the event that any payment or benefit by the Company would be subject to excise tax, then the cash severance payments shall be reduced. All payments and benefits by the Company shall hereinafter be referred to as the "Total Payments." This includes the payments and benefits under Section 3(a) hereof."

The researchers found that when they tested people on their ability to understand and recall the meaning of a legal text, their performance improved the most when center-embedded structures were replaced with more straightforward sentences, with terms defined separately.

"Using center-embedded clauses is standard writing practice in legal documents, and it makes the text very difficult to understand. It's memory-intensive for anyone, including lawyers," Gibson says. "This is something you could change and not affect the meaning in any way, but improve the transmission of the meaning."

Another feature that contributed to the incomprehensibility of legal documents was the use of uncommon words such as "lessee" and "lessor." The researchers found that replacing these words with more common alternatives such as "tenant" and "landlord" improved readers' ability to understand and recall the meaning of what they had read.

"We found more words that could have been simplified in legal text than in any other genre that we looked at, including academic text," Martinez says.

A plea for plainer language

One argument that legal theorists have put forth for why legal documents are written the way they are is that the language needs to be complex in order to unambiguously convey the meaning of complicated related concepts. However, the MIT researchers believe that is not true, because they found that many of the jargon terms used in legal documents can be replaced with more common words without changing the meaning, and the center-embedded clauses can also be replaced by nonembedded clauses to give rise to the same meaning.

Another possibility that the MIT researchers raise is that lawyers may not want to change the way they write, either because it's what they're used to or because they want their documents to look "professional" and be taken more seriously by their
Efforts to write legal documents in plainer language date to at least the 1970s, when President Richard Nixon declared that federal regulations should be written in "layman's terms." However, another study by Martinez, Mollica, and Gibson, not yet published, suggests that legal language has changed very little since that time. The researchers hope that their *Cognition* study, which points out specific aspects of legal language that make it more difficult to understand, will help inspire those who write legal documents to make a greater effort to improve the clarity of their documents.

"This is the first time someone has been able to say, 'Here's what makes legal language difficult to understand.' Before, they just speculated, and maybe that's why it hasn't changed," Gibson says. "If people know what makes it hard to understand, then maybe they can work on fixing it."

Making legal documents easier to understand could help anyone who needs to read such documents, but would have the most benefit for people who are unable to hire lawyers to help them, the researchers say.

"This is something that is especially important for those who are not able to afford legal counsel to help them understand the law," Martinez says. "If you can't afford to hire an attorney, then being able to read the documents on your own will better equip you to understand your rights."
