This corpus study investigates how young children figure out the basic clause types of their language (*declaratives*, *interrogatives*, and *imperatives*) by examining the kinds of clause types and speech acts children are exposed to in the first years of life, using data from the Providence portion of CHILDES. Early knowledge of clause types would be useful for learners acquiring the basic syntactic dependencies of their language, such as word order, argument structure, or wh-movement. The acquisition of declarative clauses would be particularly useful, as they are 'basic,' in some sense (Pinker 1984) and would likely serve as a catalyst for identifying syntactic displacement phenomena (Perkins 2019).

But how could children figure out the clause types before figuring out the phrasal dependencies? We hypothesize that children learn the clause types via the illocutionary force of the utterance. The cognitive development literature shows that young children are sensitive to the intentions and goals of speakers around them, and in particular to the intended force of their illocutionary acts (Shatz 1978, Spekman & Roth 1985, Grosse & Tomassello 2012). This opens the possibility that even children who do not yet have the grammatical knowledge to fully parse their input nonetheless understand something about what an adult is trying to accomplish with their utterances: proffer a belief, solicit an answer, make a request. If there is regularity in the observed communicative act and the morphosyntactic properties of the clause used, children could use these regularities to link particular clusters of morphosyntactic properties with particular kinds of acts. These associations would then provide a basis for sorting the other clause types and identify syntactic dependencies like argument structure or wh-movement. This hypothesis depends on the existence of a relatively stable link between speech act and the type of clause used; that is, there needs to be relatively few *indirect* speech acts in the input. Indirect speech acts, for example, would be when an interrogative clause makes a request (e.g., 'can you pass the salt?) Furthermore, it would be useful if mismatches between speech act and typical clause type were somehow marked.

In this study we ask whether children's input displays regularities in the association between clause type and speech act, and whether this aspect of the input changes over time. Second, and central to our point, we ask how regularly clause types and speech act 'match,' i.e., how often assertions are made with declaratives, questions with interrogatives, and requests with imperatives. Lastly, we ask if there are systematic patterns in the mismatches which would aid the child in flagging them such that they don't constitute counter-examples to the potential generalizations.

Two annotators marked ~500 adult utterances from 32 different transcripts of recorded play-time sessions between 6 parent/child pairs between ages 1;0 and 3;07 -- 12 transcripts within age group 1 (2 per child), 12 transcripts within age group 2 (2 per child), and 8 transcripts at within age group 3 (at least 1 transcript for 5 of the children; one child was not recorded past the age of 2). Each utterance was categorized by form: *declarative* (tensed clause), *interrogative* (Wh; subject-auxiliary-inversion), *imperative* (uninflected VP). The accompanying audio/visual context was evaluated to determine the intended speech act: *assertion*, *question*, *order/request*. We understand that there are more speech acts than these three — however, the principles that guided the coding for each speech act were relatively broad such that they could capture the basic effect of each clause type. For an assertion, we'd ask, is the mother committing to the belief that *p* (an *assertion*); is the mother committing to a preference that the child does *p* (a *request/order*); is the mother soliciting an answer that pertains to *p* (a *question*)? In addition, both annotators marked up two of the same transcripts and had an agreement rate of 97.8%.

In our study, we find that declarative clauses are the most numerous clause overall, making up 58.4% of the clause type data, interrogatives occur at a rate of 25%, and imperatives

are the least used at 16.5%. When we break down by age, we see that this general hierarchy is maintained. There are slight differences: interrogatives are used in 30% of the data for 1 year olds, but drops to 22% and 21% in 2 year olds and 3 year olds, respectively.

Speech act and clause type across ages

Age ↓	Clause type → Speech act ↓	Declarative	%	Imperative	%	Interrogative	%	Total
1	Assertion	1566	94.5%	75	4.6%	5	0.3%	1646
	Question	204	17.3%	6	0.5%	970	82.2%	1180
	Request	95	13.7%	516	74.4%	83	12.0%	694
2	Assertion	2129	95.4%	97	4.3%	6	0.3%	2232
	Question	238	22.3%	11	1.0%	816	76.6%	1065
	Request	128	18.7%	470	68.6%	87	12.7%	685
3	Assertion	1314	97.8%	17	1.3%	13	1.0%	1344
	Question	146	22.1%	4	0.6%	510	77.3%	660
	Request	95	16.0%	459	77.5%	38	6.4%	592

Figure 1: This salmon-colored boxes highlight the percentage that each speech act was used with its 'canonical' clause type, i.e., the percentage assertions were made with declaratives, requests with imperatives, and questions with interrogatives.

Assertions are very tightly correlated with declaratives; across the ages, about 96% of them were made with declaratives. Questions and requests are much more variable in terms of what kinds of clauses can express them. Questions were made with declaratives 20% of the time after averaging across the ages. Declarative questions are all instances of *rising declaratives*, and thus, are marked by prosody. In total, 16% of the requests were made with declaratives, and 10% were made with interrogatives. We find a pattern, however, that most declarative requests have a modal in them: 59.7% (e.g., 'no you have to leave that on honey'), and interrogative requests have a modal in them 73% of the time (e.g., 'can you move your foot please?'). The mismatches that occur the most, thus, have elements that could be tracked by the child and thus used to explain away possible counterexamples to the clause type-speech act mapping generalizations.

There is more work to be done to fully test the hypothesis that children use illocutionary force to identify the clause types of their language. It would have to be true that (i) there is a consistent signal in speech to children reflecting mapping between speech act and clause type (ii) children are sensitive to the elements involved in the mapping (iii) children use the correlations between form and speech act to learn. We now have evidence to support step (i).

References: **Grosse**, G. **& Tomasello**, M. (**2012**). Two-year-old children differentiate test questions from genuine questions. Journal of Child Language. **Perkins**, L. (**2019**). How Grammars Grow: Argument Structure and the Acquisition of Non-Basic Syntax. University of Maryland dissertation. **Pinker**, S. (**1984**). Language learnability and language development. Cambridge, MA Harvard University Press. **Shatz**, M. (**1978**). On the development of communicative understandings: An early strategy for interpreting and responding to messages. Cognitive Psychology. **Spekman**, N. J. **& Roth**, F. P. (**1985**). Preschool children's comprehension and production of directive forms. Journal of Psycholinguistic Research.