

Rapid Acquisition of Disjunction from Prosody and Consistency

Masoud Jasbi

masoudj@stanford.edu
Department of Linguistics
Stanford University

Akshay Jaggi

ajaggi@stanford.edu
Department of Linguistics
Stanford University

Michael C. Frank

mcfrank@stanford.edu
Department of Psychology
Stanford University

Abstract

You need to mention the conflict between the usage-based accounts and the nativist accounts. You need to say that the exclusivity of disjunction can be learned from prosody and inconsistency of the disjuncts themselves (e.g. dirty or clean). Consistent with the nativist account, we find that even though *or* is rare in child-directed speech compared to *and*, it is learned relatively quickly by children and produced at an adult rate by the age four. We also show that consistent with usage-based accounts of learning, there is a consistent pattern between usage and interpretation that children can learn from.

Keywords: language acquisition; word learning; logical words; and; or; nativism; constructivism.

Introduction

“James Bond producer says next 007 could be black **or** a woman”, reported the social media and entertainment company LADbible in a tweet. A twitter user named Robert responded sarcastically with: “if only women could be black!” What in the producer’s speech gave the impression that the next 007 could not be both black **and** a woman? The word *or* is associated with two interpretations: **inclusive**, and **exclusive**. A disjunction like “A or B” is inclusive when it is interpreted as “A or B or both”. This is probably what LADbible meant when reporting the James Bond producer. However, “A or B” can also be interpreted as exclusive: “A or B but **not both**”. Robert’s response shows that he had an exclusive interpretation of *or*. What factors determine the interpretation of *or* and how do children learn its meaning given this ambiguity?

A large body of research in linguistics and philosophy in the past 50 years has created a common consensus on the meaning of *or* (see Aloni (2016)). Data on the interpretation of disjunction across different sentences, contexts, and even languages show that the core meaning of disjunction words such as *or* is **inclusive**. This is similar to the definition of disjunction in formal logics. The **exclusive** interpretation of *or* is the result of enhancing its inclusive semantics via other (extra-semantic) factors such as intonation (Pruitt & Roelofsen, 2013), inconsistency of the options (Geurts, 2006), and pragmatic reasoning over the speaker’s choice of *or* instead of *and* (Grice, 1975). In other words, interpreting a disjunction is a complex process that needs to take into account the meaning of *or* as well as different structural and contextual factors that accompany it. How do we, as children, learn such a complex interpretive system?

There are two accounts of children’s acquisition of disjunction: a constructivist account, and a nativist account. Under the constructivist account, children learn the meaning of

or by paying attention to how parents use it in different contexts. They form usage-rules and expand their usage repertoire as they grow up. The prediction is that children’s production of *or* in children mirrors what they hear from parents. Morris (2008) found that *or* is much less frequent than *and* in parents’ speech to children. As predicted by the constructivist account, children also learn to produce *and* quickly while for *or* there is a gradual increase in production, possibly reaching the adult level at age five or six. In children’s speech, *and* appears quickly and around the age 2 or 3 while it takes several years for *or* to appear. The faster acquisition of *and* is consistent with the constructivist theory that emphasizes frequency of usage in children’s linguistic development. Morris also showed that the majority of *or* examples children hear are exclusive. He showed that consistent with the constructivist account, the majority of *ors* children produce are also exclusive.

However, several comprehension studies of *or* in different linguistic contexts show that children between the ages of three and five interpret *or* as inclusive disjunction (Crain, 2012). This is a surprising finding given the previous finding that the majority of *or* examples children hear are exclusive. Crain (2012) suggests that children rely on innate knowledge that the disjunction operator in their native language must have an inclusive meaning. More broadly, logical nativism argues that children have an innate knowledge that the connective words in their native language should refer to operators similar to those in first order logic. Exclusive disjunction is not an operator available

Here we provide alternative answers to two puzzles raised in previous literature: 1. Production-Comprehension Discrepancy: the production data suggests that learning *or* is a gradual and long process, possibly reaching an adult-like level at the age five or six. However, comprehension studies suggest that children have an adult-like understanding of *or* at the age of four or even three. 2. Exclusivity puzzle: even though the majority of *or* examples children hear are exclusive, comprehension studies show that children understand *or* as inclusive disjunction. How do children learn what they do not seem to hear?

The first study shows that despite (relatively) rare occurrences of *or*, it is learned relatively quickly and between the ages of 2 and 4. The finding confirms the nativist

Implications for future work.

Study 1: Corpus Study

First, we conducted an exploratory and large scale investigation of *and* and *or* productions in parents and children. The

goal of the study was to find out when children start producing these words and when they reach the adult rate of production. We conclude that children start producing *and* around 1.5 or 2 years of age, and *or* between the ages of 2 and 3. They reach the adult rate of production for *and* around 3 and for *or* around 4 or possibly earlier.

Methods

We accessed the Child Language Data Exchange System (CHILDES, MacWhinney (2000)) via the online platform *childes-db* and its associated R package *childesr* (Sanchez et al., in prep). We extracted all instances of *and* and *or* from the English corpora (ENG-NA and ENG-UK). We limited our analysis to the data between one and six years because there is scarce data outside this age range. We computed the relative frequency of connective production by dividing the total number of *and/or* in the speech of fathers, mothers, and children at a particular age by the total number of words spoken at that age. We present the relative frequency as parts per thousand.

Results

In figure 1, we show the relative frequencies of *and* and *or* in the speech of parents and children between one and six years. It is important to note that the y-axes for *and* vs. *or* show different ranges. This is due to the big difference in the relative frequencies of *and* and *or*. In the speech of parents, *and* is produced around 20 times per thousand words while *or* is only produced around 2 times per thousand words. This confirms previous findings that *or* is much less frequent in child directed speech than a similar function word such as *and*.

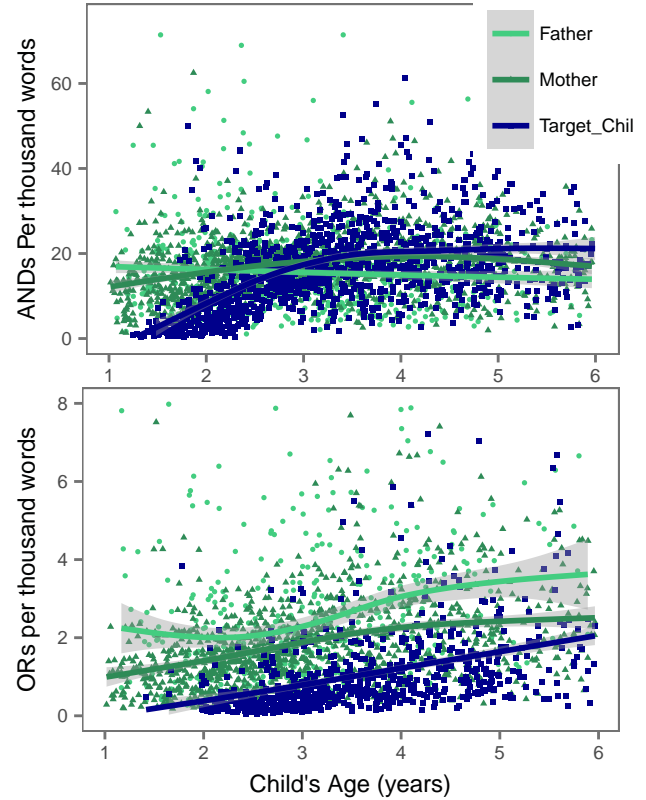


Figure 1: The relative frequency of AND (top) and OR (bottom) per thousand words in the speech of fathers, mothers, and children between the ages of 1 and 6.

And and *or* seem to show different developmental trajectories in the speech of children. For *and*, there is a rapid increase in its production between the ages of 1.5 and 3 before it reaches the adult rate around the age 3 and stay at that level until the age 6. For *or*, on the other hand, we see a slow increase from the age 2 until the age 6 when it reaches the adult rate. This difference in the development of *and* & *or* production was attributed to the frequency of these items in child-directed speech. Since *and* is much more frequent than *or*, it is learned much faster than *or*. Morris (2008) argued that such patterns support the item-based and usage-based acquisition of logical words.

However, the analysis above does not control for other factors that can affect the production of words by children. An important factor to control for is the development of speech acts. While content words such as *dog* or *chair* may appear freely in different types of speech acts, function words are highly constrained by the type of speech acts they can appear in. For example, it is reasonable to assume that question words such as *how* and *why* are much more likely to occur in questions than statements (declaratives). If parent-child interaction is such that parents ask more questions than children, it is not surprising to find higher rates of *how* and *why* production in parents than children. Therefore, it is important to control for the speech act a function word appears in.

Figure 2 shows the relative frequencies of *and* and *or* in questions vs. declaratives, in the speech of parents and children between one and six years. Here, the relative frequency is computed by dividing the total number of *and/or* in a question/declarative in the speech of fathers, mothers, and children at a particular age, by the total number of words in a question/declarative spoken at that age. As before, we present the relative frequency as parts per thousand.

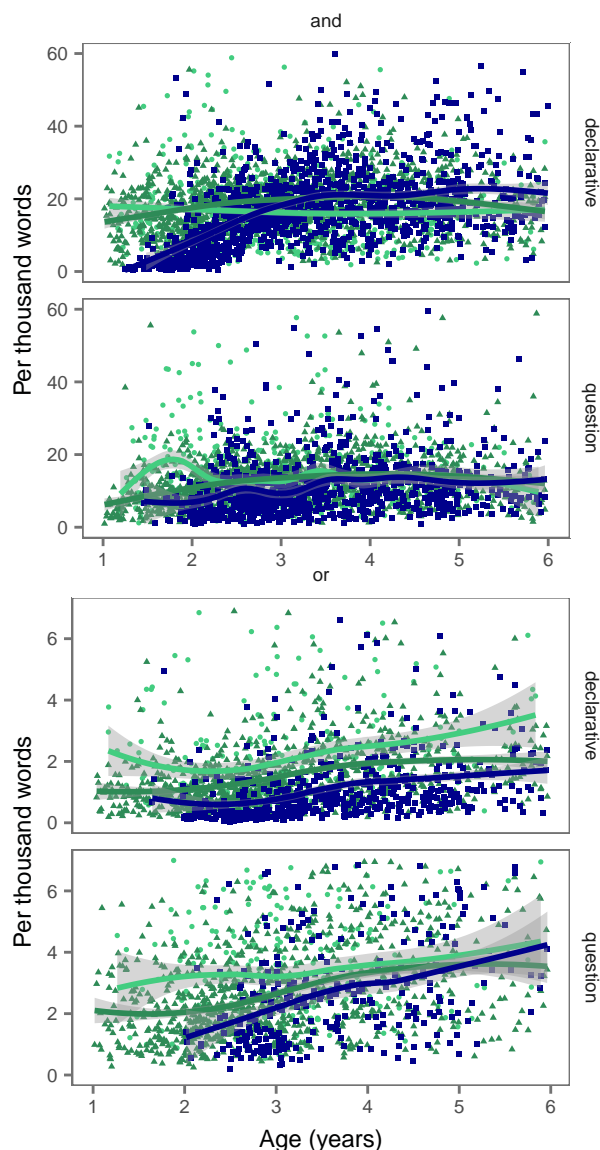


Figure 2: The relative frequency of AND (top) and OR (bottom) per thousand words in declaratives and questions in the speech of fathers, mothers, and children between the ages of 1 and 6,

The results show similar developmental trajectories for the production of *and* and *or* in children. For both words, there is a relatively rapid increase in their frequency between the ages of 2 and 4 before reaching the parent rate at the age of 4 and staying at that rate until the age of 6. This pattern of

production is consistent with the nativist observations that the acquisition of *and* and *or* is rapid and that children have an adult-like comprehension of these two connectives at the age 4.

Summary

And is a lot more frequent than or in child directed speech.

In the first six years, it appears that children reach the adult production rate for *and* but not *or*.

This is at least partly because *or* is more frequent in questions and children produce fewer questions than parents.

The developmental trajectory of connective production is best described as a quick increase in production between the ages of 2 and 4 and staying around the parents' rate between the ages 4 and 6. This is compatible with the comprehension studies which suggest children understand *and* and *or* by the age four.

Study 2: Annotation Study

Use standard APA citation format. Citations within the text should include the author's last name and year. If the authors' names are included in the sentence, place only the year in parentheses, as in (???), but otherwise place the entire reference in parentheses with the authors and year separated by a comma (???). List multiple references alphabetically and separate them by semicolons (???, ???). Use the et. al. construction only after listing all the authors to a publication in an earlier reference and for citations with four or more authors.

For more information on citations in R Markdown, see [here](#).

Footnotes

Indicate footnotes with a number¹ in the text. Place the footnotes in 9 point type at the bottom of the page on which they appear. Precede the footnote with a horizontal rule.²

Figures

All artwork must be very dark for purposes of reproduction and should not be hand drawn. Number figures sequentially, placing the figure number and caption, in 10 point, after the figure with one line space above the caption and one line space below it. If necessary, leave extra white space at the bottom of the page to avoid splitting the figure and figure caption. You may float figures to the top or bottom of a column, or set wide figures across both columns.

Two-column images

You can read local images using png package for example and plot it like a regular plot using grid.raster from the grid package. With this method you have full control of the size of your image. **Note: Image must be in .png file format for the readPNG function to work.**

You might want to display a wide figure across both columns. To do this, you change the `fig.env` chunk option

¹Sample of the first footnote.

²Sample of the second footnote.

to figure*. To align the image in the center of the page, set fig.align option to center. To format the width of your caption text, you set the num.cols.cap option to 2.

One-column images

Single column is the default option, but if you want set it explicitly, set fig.env to figure. Notice that the num.cols option for the caption width is set to 1.

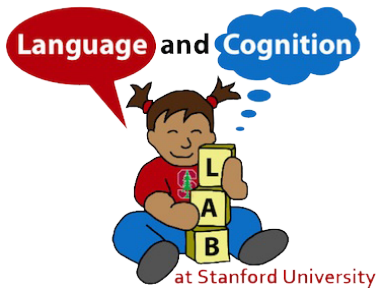


Figure 4: One column image.

R Plots

You can use R chunks directly to plot graphs. And you can use latex floats in the fig.pos chunk option to have more control over the location of your plot on the page. For more information on latex placement specifiers see [here](#)

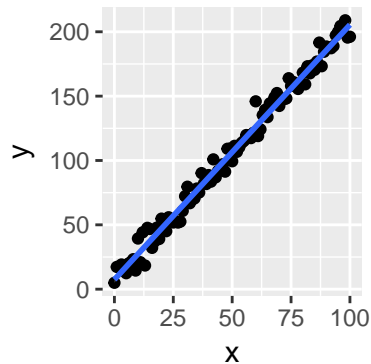


Figure 5: R plot

Tables

Number tables consecutively; place the table number and title (in 10 point) above the table with one line space above the caption and one line space below it, as in Table 1. You may float tables to the top or bottom of a column, set wide tables across both columns.

You can use the xtable function in the xtable package.

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.06	0.11	0.6	0.57
x	1.95	0.11	18.5	0.00

Table 1: This table prints across one column.

Conclusion (Masoud)

- List what you showed and argued for.
- Talk about future directions.
- Possibly talk about the original tweet and how that falls still outside what your algorithm learns. # Acknowledgements
- Place acknowledgments (including funding information) in a section at the end of the paper.

References

Aloni, M. (2016). Disjunction. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy*. Stanford University. Retrieved from <https://plato.stanford.edu/archives/win2016/entries/disjunction/>

Crain, S. (2012). *The emergence of meaning*. Cambridge University Press.

Geurts, B. (2006). Exclusive disjunction without implications. *Ms., University of Nijmegen*.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3: Speech Acts, pp. 43–58). Academic Press.

MacWhinney, B. (2000). *The chldes project: The database* (Vol. 2). Psychology Press.

Morris, B. J. (2008). Logically speaking: Evidence for item-based acquisition of the connectives and & or. *Journal of Cognition and Development*, 9(1), 67–88.

Pruitt, K., & Roelofsen, F. (2013). The interpretation of prosody in disjunctive questions. *Linguistic Inquiry*, 44(4), 632–650.

Sanchez, A., Meylan, S., Braginsky, M., MacDonald, K., Yurovsky, D., & Frank, M. C. (in prep). Childes-db: A flexible and reproducible interface to the child language data exchange system (childes).



Figure 3: This image spans both columns. And the caption text is limited to 0.8 of the width of the document.