Understanding and Awareness Data Deposit information

This deposit contains the diverse experimental and meta-analytic datasets collected during the Understanding and Awareness project. The datasets assess psychological research questions involving the relationship between how we understand and use language, and how we attend to the world around us. For example, one dataset investigates whether words and sentences can be partially understood when they have been masked from conscious awareness. Another dataset investigates how preschool children allocate their attention when describing scenes that require them to use potentially ambiguous language.

Note that this project did not collect one large dataset, but rather a range of different datasets, with many different characteristics; fuller descriptions of each dataset are provided in the uploaded documentation file.

The datasets in this deposit report 1) Chronometric (response time) studies conducted with adults and with preschool children (aged 3 and 5). 2) Eye tracking studies conducted with adults and with preschool children (aged 2 through 5). 3) Psychophysical (continuous flash suppression) studies conducted with adults. 4) Looking time experiments conducted with infants (age 7 months). 5) A database containing records of a meta-analysis of infant looking time data.

The size of the datasets range from a meta-analysis containing approximately 100 records, to a collection of psychophysical datasets containing thousands of records for 100s of subjects.

The datasets described in this deposit are not stored on UK DataStore, rather, following UK DataStore's request, this document provides links to repositories for the data that are held on either the Open Science Foundation's website, or on GitHub.

The remainder of this document provides more details on each of the repositories. Each repository is linked to a published paper; links to the repositories are included in the detailed discussion later in this document, as well as in the references below.

Page 3. Rabagliati, H., Ferguson, B., & Lew-Williams, C. (in press). The profile of abstract rule learning in infancy. Meta-analytic and experimental evidence. Developmental Science.
DOI: **10.1111/desc.12704**
Repository: https://osf.io/5k3vw/

Page 4. Rabagliati, H. , Robertson, A., & Carmel, D. (2018). The importance of awareness for understanding language. Journal of Experimental Psychology: General, 147(2), 190-208.
DOI: 10.1037/xge0000348.
Repository: https://github.com/hughrabagliati/CFS_Compositionality

Page 6. Rabagliati, H.A., Doumas, L.A.A., & Bemis, D.K. (2017). Representing composed meanings through temporal binding. Cognition, 162, 61-72.
DOI: 10.1016/j.cognition.2017.01.013
Repository: https://github.com/hughrabagliati/CompositionalityPaper1

Page 8. Rabagliati, H., & Robertson, A. (2017). How do children learn to avoid referential ambiguity? Insights from eye-tracking. Journal of Memory and Language, 94, 15-27.
DOI: 10.1016/j.jml.2016.09.007
Repository: **https://github.com/hughrabagliati/ETRef**

Page 10. Gambi, C., Pickering, M.J., & Rabagliati, H. (2016). Beyond associations: Sensitivity to structure in pre-schoolers' linguistic predictions. Cognition, 157, 340-351.
DOI: 10.1016/j.cognition.2016.10.003
Repository: **https://github.com/chiara-gambi/structpred.**

Page 12. Lindsay, L., Gambi, C., & Rabagliati, H. (in press). Preschoolers optimize the timing of their conversational turns through flexible coordination of language comprehension and production. Psychological Science.
DOI: 10.1177/0956797618822802
Repository: https://osf.io/kcp9z/

Details of deposit.

1. Meta-analytic and experimental evidence about infant pattern learning.
Rabagliati, H., Ferguson, B., & Lew-Williams, C. (in press). The profile of abstract rule learning in infancy. Meta-analytic and experimental evidence. Developmental Science. DOI: **10.1111/desc.12704**

Everyone agrees that infants possess general mechanisms for learning about the world, but the existence and operation of more specialized mechanisms is controversial. One mechanism—rule learning—has been proposed as potentially specific to speech, based on findings that 7-month-olds can learn abstract repetition rules from spoken syllables (e.g. ABB patterns: wo-fe-fe, ga-tu-tu…) but not from closely matched stimuli, such as tones. Subsequent work has shown that learning of abstract patterns is not simply specific to speech. However, we still lack a parsimonious explanation to tie together the diverse, messy, and occasionally contradictory findings in that literature. We took two routes to creating a new profile of rule learning: meta-analysis of 20 prior reports on infants' learning of abstract repetition rules (including 1,318 infants in 63 experiments total), and an experiment on learning of such rules from a natural, non-speech communicative signal. These complementary approaches revealed that infants were most likely to learn abstract patterns from meaningful stimuli. We argue that the ability to detect and generalize simple patterns supports learning across domains in infancy but chiefly when the signal is meaningfully relevant to infants' experience with sounds, objects, language, and people.

This project investigates infant abstract rule learning (Marcus et al, 1999) and its potential domain specificity (e.g., Marcus et al., 2007; Saffran et al., 2007; Ferguson & Lew-Williams, 2016). The project incorporates a meta-analysis of the infant rule learning literature, and a novel experiment.

For the meta-analysis, the OSF Storage component (https://osf.io/5k3vw/) of this repository contains an R Markdown document that provides full details of the analyses carried out, as well as associated graphics such as a Forest plot and PRISMA flowchart. The associated Google Drive storage contains the raw data used in the meta-analysis.

For the experiment the OSF storage component of this repository (https://osf.io/5k3vw/) contains an R Markdown document providing the full data, processing and analysis protocols, as well as all of the experimental stimuli and materials and raw data.

2. Psychophysical studies of the relation between attention, consciousness and language processing

Rabagliati, H. , Robertson, A., & Carmel, D. (2018). The importance of awareness for understanding language. Journal of Experimental Psychology: General, 147(2), 190-208. DOI: 10.1037/xge0000348.

Is consciousness required for high level cognitive processes, or can the unconscious mind perform tasks that are as complex and difficult as, for example, understanding a sentence? Recent work has argued that, yes, the unconscious mind can: Sklar et al. (2012) found that sentences, masked from consciousness using the technique of continuous flash suppression (CFS), broke into awareness more rapidly when their meanings were more unusual or more emotionally negative, even though processing the sentences' meaning required unconsciously combining each word's meaning. This has motivated the important claim that consciousness plays little-to-no functional role in high-level cognitive operations. Here, we aimed to replicate and extend these findings, but instead, across 10 high-powered studies, we found no evidence that the meaning of a phrase or word could be understood without awareness. We did, however, consistently find evidence that low-level perceptual features, such as sentence length and familiarity of alphabet, could be processed unconsciously. Our null findings for sentence processing are corroborated by a meta-analysis that aggregates our studies with the prior literature. We offer a potential explanation for prior positive results through a set of computational simulations, which show how the distributional characteristics of this type of CFS data, in particular its skew and heavy tail, can cause an elevated level of false positive results when common data exclusion criteria are applied. Our findings thus have practical implication for analyzing such data. More importantly, they suggest that consciousness may well be required for high-level cognitive tasks such as understanding language.

This repository contains psychophysical studies on whether word meanings are combined into sentence meanings when they are masked from awareness using the technique of Continuous Flash Suppression. The reported experiments are based on those reported in Sklar et al. (2012). Reading and doing arithmetic nonconsciously. PNAS. The key finding of these studies is that we could not replicate the claims of Sklar and colleagues, such that we found no evidence that participants could reach nonconsciously.

Data from these studies can be found in the GitHun repository at https://github.com/hughrabagliati/CFS_Compositionality

Details of the data and the analyses performed can be found in the markdown file at https://github.com/hughrabagliati/CFS_Compositionality/blob/master/Overall_Analysis.Rmd. For more casual readers, an interpreted version of that Markdown file can be found here.

The deposit contains data from four major experiments as well as from a set of statistical simulations:

Experiment 1

Experiment 1a replicates Experiment 1 of Sklar et al (2012), but with English stimuli. That experiment demonstrated lower suppression times for sentence with semantically unusual meanings. Experiment 1b is an extension in which we compare sentences to their reverse (e.g., Tom whisked the cream, The cream whisked Tom), in order to better control lexical statistics.

Experiment 2

Experiment 2a replicates Expriment 4 of Sklar et al (2012), but with English stimuli. That experiment demonstrated that phrases with negative emotional meanings (concentration camp) break suppression faster than phrases with neutral meanings. Experiment 2b is an extension in which we compare sentences to their reverse (the baby hit the brick, the brick hit the baby). Experiment 2c is a replication of Sklar et al.'s original Hebrew experiment, to test whether visual properties of Hebrew influence participants suppression times.

Experiment 3

This replicated our Experiments 1a, 1b, 2a and 2b, but using the original presentation scripts used in Sklar et al. (2012).

Experiment 4

This replicated a single-word emotional CFS experiment, by Yang and Yeh (2011, Consciousness and Cognition).

Simulations

The file sensicality_sims.r details a series of simulations that we ran, investigating how skewed data (of the type found in breaking continuous flash suppression experiments) can affect parametric statistical analyses such as t-tests.

3. Response time studies of how adults combine word meanings to form sentence meanings

Rabagliati, H.A., Doumas, L.A.A., & Bemis, D.K. (2017). Representing composed meanings through temporal binding. Cognition, 162, 61-72. DOI: 10.1016/j.cognition.2017.01.013

A key feature of human thought and language is compositionality, the ability to bind pre-existing concepts or word meanings together in order to express new ideas. Here we ask how newly composed complex concepts are mentally represented and matched to the outside world, by testing whether it is harder to verify if a picture matches the meaning of a phrase, like big pink tree, than the meaning of a single word, like tree. Five sentence-picture verification experiments provide evidence that, in fact, the meaning of a phrase can often be checked just as fast as the meaning of one single word (and sometimes faster), indicating that the phrase's constituent concepts can be represented and checked in parallel. However, verification times were increased when matched phrases had more complex modification structures, indicating that it is costly to represent structural relations between constituent concepts. This pattern of data can be well-explained if concepts are composed together using two different mechanisms, binding by synchrony and binding by asynchrony, which have been suggested as solutions to the "binding problem" faced in both vision science and higher-level cognition. Our results suggest that they can also explain aspects of compositional language processing.

This deposit contains the results of five response time experiments, conducted over the internet. It can be found at https://github.com/hughrabagliati/CompositionalityPaper1 .

In these studies, adult participants made sentence-picture verification judgments; for example, they would say a phrase, like “red boat”, and then have to decide if a subsequent picture showed a red boat. We varied the type of phrase, and the task participants performed. For example, sometimes participants saw a full phrase, and sometimes just a single word (e.g., “boat”). In some tasks, participants had to judge whether both words matched the picture (eg., is the picture a red boat?) and in other tasks participants had to judge whether one of the words matched the picture (e.g., is the picture red or a boat?).

The deposit is made up of a series of folders and R scripts. In each folder is the anonymized individual participant data from each study. The matched R scripts will read in the data, perform statistical analysis, and produce the figures reported in the paper.

Experiments 1a, 1b, 1c

These report studies comparing verification judgments about single words (e.g., “boat”) to judgments about two word phrases (e.g., “red boat”).

Experiments 2 and 3

These report studies comparing verification judgments about single words, two-word phrases, and three-word phrases, in order to test the specific predictions of a computational model of how word meanings are combined.

4. Eye tracking studies of how children and adults allocate attention when using language to describe the world

Rabagliati, H., & Robertson, A. (2017). How do children learn to avoid referential ambiguity? Insights from eye-tracking. Journal of Memory and Language, 94, 15-27. DOI: 10.1016/j.jml.2016.09.007

Children have considerable difficulty producing informative and unambiguous referring expressions, a fact that still lacks a full explanation. Potential insight can come from psycholinguistic models of ambiguity avoidance in adults, which suggest that, before describing any scene, speakers pro-actively monitor for some — but not all — types of potential ambiguity, and then subsequently monitor whether their just-produced expression provides an ambiguous description. Our experiments used eye tracking to assess the developing roles of these skills in children's referential communication. Experiment 1 shows that adults' eye movements can index the processes of both pro-active and self-monitoring. Experiments 2 and 3 show that children (n = 110) typically do not pro-actively monitor for potential ambiguity, although they do show evidence of pro-active monitoring on the occasions when they produce informative expressions. However, we do find evidence that children consistently monitor their own descriptions for ambiguity, even though they rarely correct their utterances. We propose that the process of self-monitoring might act as a learning signal, that guides children as they acquire the ability to monitor pro-actively.

This repository contains data from eye tracking studies of referential communication games, which are simple tasks in which preschoolers and adults had to identify simple target pictures to a confederate in the presence of ambiguity (e.g., if asked to identify a depicted penguin in the presence of another type of penguin, you would to specify which penguin you were talking about).

The repository can be found at https://github.com/hughrabagliati/ETRef . It contains anonymized coded data (i.e., no audio files, no names).

For all studies, except Experiment 3 which is not reported in the paper, the analysis scripts are embedded in the R Markdown Document "RefCommPaper.md" which can be found at https://github.com/hughrabagliati/ETRef/blob/master/RefCommPaper.Rmd.

Experiment 1 assesses adults' utterances and eye movements while completing a referential communication task, in which ambiguity is created by either semantic or phonological overlap. **Experiment 1a** is a silent version of this task, in which adults simply communicated by pointing.

Experiment 2 assesses 4-to-5-year-olds' utterances and eye movements while completing a referential communication task, in which ambiguity is created by semantic overlap.

Experiment 3 is not yet added, but assesses 4-to-5-year-olds' performance and eye movements on an explicit ambiguity detection task. The task is not reported in the paper (but the kids were pretty good at it).

Experiment 4 assesses 4-to-5-year-olds' performance on a referential ambiguity task where they must name both ambiguous items. It therefore tests whether young children use production based monitoring.

5. Eye tracking studies of children's semantic predictions during sentence processing

Gambi, C., Pickering, M.J., & Rabagliati, H. (2016). Beyond associations: Sensitivity to structure in pre-schoolers' linguistic predictions. Cognition, 157, 340-351. DOI: 10.1016/j.cognition.2016.10.003

One influential view of language acquisition is that children master structural generalizations by making and learning from structure-informed predictions. Previous work has shown that from 3 years of age children can use semantic associations to generate predictions. However, it is unknown whether they can generate predictions by combining these associations with knowledge of linguistic structure. We recorded the eye movements of pre-schoolers while they listened to sentences such as *Pingu will ride the horse.* Upon hearing *ride,* children predictively looked at a horse (a strongly associated and plausible patient of *ride*), and mostly ignored a cowboy (equally strongly associated, but an implausible patient). In a separate experiment, children did not rapidly look at the horse when they heard *You can show Pingu … "riding"*, showing that they do not quickly activate strongly associated patients when there are no structural constraints. Our findings demonstrate that young children's predictions are sensitive to structure, providing support for predictive-learning models of language acquisition.

This deposit contains data files for two studies that assessed how children allocate attention around the world while they are listening to meaningful sentences. In particular, it tests whether children attend to external information whose mention is predictable given the semantic and syntactic structure of the sentence that they are hearing.

The deposit can be found at https://github.com/chiara-gambi/structpred.

We conducted two experiments as part of this project. In Experiment 1, we tested both adults and children (3-to-5 yo). In Experiment 2, we tested only children (3-to-5 yo). Please refer to the published article for details on the number of participants and their demographics. Metadata for the datafiles can be found in the depost.

Both experiments used eye-tracking while listening. In Experiment 1, participants heard a spoken sentence that contained either a Predictive or a Non-Predictive verb (this is the first factor used in our analyses). Sentences could also be Long or Short. Short sentences were simple active transitive sentences such as "Pingu will ride/pull the horse". Long sentences had additional modifiers inserted before the final noun (e.g., "Pingu will ride/pull the very tired horse"). This was the second factor in our analyses. While listening, participants observed 4 entities: Subject (always the well-known cartoon character Pingu), Agent (a prototypical agent of the predictive verb; e.g., cowboy), Patient (a protoypical patient of the predictive verb; e.g., horse), and a Distractor. In Experiment 2, participants listened to sentences like "Now, you can show Pingu...riding/pulling!", and Verb Type (Predictive/Non-Predictive was the only factor).

For each group of participants, we conducted two types of analyses: Snapshot analyses and Growth Curve Analyses.

Snapshot Analyses. In Experiment 1, these analyses included both the factor Length (Long/Short) and the factor Verb Type (Predictive/Non-Predictive). They compared the likelihood of fixations to the Patient (or Agent) across conditions. They were restricted to a window ranging from 200 ms before to 100 ms after the onset of the critical final word of the sentence. In Experiment 2, the time window ranged from 200 ms before to 100 ms after the offset of the sentence final verb, and Verb Type was the only factor.

Growth Curve Analyses. These analyses only included the factor Verb Type and compared fixation proportions to the Patiemt (or Agent) across conditions over time. They were restricted to a window ranging between 500 ms before and 1700 ms after the offset of the critical verb (Experiment 1), or between 500 ms before and 200 ms after verb offset (Experiment 2).

6. Chronometric analysis of preschool children's linguistic turn-taking

Lindsay, L., Gambi, C., & Rabagliati, H. (in press). Preschoolers optimize the timing of their conversational turns through flexible coordination of language comprehension and production. Psychological Science. DOI: 10.1177/0956797618822802

Conversation is the natural setting for language learning and use, and a key property of conversation is the smooth taking of turns. In adult conversations, delays between turns are minimal (typically 200ms or less) because listeners display a striking ability to predict what their partner will say, and they formulate a response before their partner's turn ends. Here, we tested how this ability to coordinate comprehension and production develops in preschool children. In an interactive paradigm, 106 children (ages 3–5 years) and 48 adults responded to questions that varied in predictability but were controlled for linguistic complexity. Using a novel distributional approach to data analysis, we found that when children can predict a question's ending, they leave shorter gaps before responding, suggesting that they can optimize the timing of their conversational turns like adults do. In line with a recent ethological theory of turn taking, this early competency helps explain how conversational contexts support language development.

This project investigates conversational turn taking in preschool children. In the reported experiment, children played an iPad game in which they guided a character around a maze. When the character reached a cross-roads, they asked the child a question about which direction they should move in. We varied the contents of that question and recorded the response time of the child.

The data and all stimulus presentation materials, along with analysis code and a preprint, can be found at https://osf.io/kcp9z/. A description of the data files and analysis code can be found at https://osf.io/j2srp/.