Phillip M. Alday*, Matthias Schlesewsky and Ina Bornkessel-Schlesewsky

Discovering Prominence and Its Role in Language Processing: An Individual (Differences) Approach

Abstract: It has been suggested that, during real time language comprehension, the human language processing system attempts to identify the argument primarily responsible for the state of affairs (the "actor") as quickly and unambiguously as possible. However, previous work on a prominence (e.g. animacy, definiteness, case marking) based heuristic for actor identification has suffered from underspecification of the relationship between different cue hierarchies. Qualitative work has yielded a partial ordering of many features (e.g. MacWhinney et al. 1984), but a precise quantification has remained elusive due to difficulties in exploring the full feature space in a particular language. Feature pairs tend to correlate strongly in individual languages for semantic-pragmatic reasons (e.g., animate arguments tend to be actors and actors tend to be morphosyntactically privileged), and it is thus difficult to create acceptable stimuli for a fully factorial design even for binary features. Moreover, the exponential function grows extremely rapidly and a fully crossed factorial design covering the entire feature space would be prohibitively long for a purely within-subjects design. Here, we demonstrate the feasibility of parameter estimation in a short experiment. We are able to estimate parameters at a single subject level for the parameters animacy, case and number. This opens the door for research into individual differences and population variation. Moreover, the framework we introduce here can be used in the field to measure more "exotic" languages and populations, even with small sample sizes. Finally, pooled single-subject results are used to reduce the number of free parameters in previous work based on the extended Argument Dependency Model (Bornkessel-Schlesewsky and Schlesewsky 2006, 2009, 2013, in press; Alday et al. 2014).

Multimedia: OpenSesame experiment and Python support scripts, sample stimuli, R scripts for analysis

Keywords: computational model, language processing, emergence, ambiguity resolution, actor identification, prominence, individual differences

DOI 10.1515/lingvan-2014-1013

1 Introduction

Parameter underspecification is a critical issue in modern linguistic models, with too many parameters typically dismissed to the periphery of qualitative description and "performance". The return on investment for working out the precise mechanistic and quantitative "details" of a model often seems too poor, especially in light of the many levels of linguistic variation: language > dialect > idiolect (inter-speaker) > intra-speaker. Yet, it is exactly these parameters and how they can vary that is interesting when discussing *language* instead of *a language*.

Even well-formulated psycholinguistic and neurolinguistic models often suffer from underspecification with many parameters omitted and many more never empirically estimated. Implemented computational models suffer less from the underspecification problem, but still have many issues with free parameters

***Corresponding author: Phillip M. Alday,** University of Marburg, 35032 Marburg, Germany, E-mail: phillip.alday@staff.uni-marburg.de
Matthias Schlesewsky, Johannes-Gutenberg University Mainz, 55128 Mainz, Germany
Ina Bornkessel-Schlesewsky, University of Marburg, 35032 Marburg, Germany; University of South Australia, GPO Box 2471, Adelaide, SA 5001, Australia

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

(Howes et al. 2009) and researcher degrees of freedom (Simmons et al. 2011). Previously, we presented a computational model of language processing based on the interaction of weighted *prominence features* (Alday et al. 2014). While our models provided a good fit for event-related potential data (i.e. electrical brain activity time-locked to a critical word within a sentence) which has a very poor signal-to-noise ratio, we nonetheless relied on a somewhat problematic conversion of ordinally scaled data to ratio-scaled data using simple logarithmic scaling. In the following we present a framework for empirically quantifying the parameters of well-defined computational models based on competition and constraint-satisfaction, focusing on the class of prominence-based models.

Using a small experiment and a basic statistical technique, we demonstrate that it is possible to estimate parameters at the single subject level in less than half an hour and perhaps a good cup of coffee. The ease of this approach opens the door to quantitative study of interindividual variation and linguistic settings in which only small samples of speakers are accessible (e.g. less-researched languages, clinical populations).

2 Prominence, the extended argument dependency model (eADM) and actor-centered comprehension

Before turning to the parameter estimation approach that is the focus of the present paper, we will briefly describe the empirical neurocognitive model on which it is based. This framework will provide two critical concepts for the parameter estimation: *prominence* (the independent variable) and the *actor role* (the dependent variable).

The extended Argument Dependency Model (eADM) is a neurocognitive, and more recently neurobiologically grounded model of cross-linguistic language comprehension which places particular emphasis on the role of the "actor" participant (Bornkessel-Schlesewsky and Schlesewsky 2006, 2009, 2013, 2014) The actor, a term taken from Role and Reference Grammar (Van Valin 2005) and termed Proto-Agent in other approaches (Dowty 1991; Primus 1999), refers to the event instigator/participant primarily responsible for the state of affairs being described. Based on the results of electrophysiological studies across a range of typologically diverse languages, the eADM posits that comprehension is actor-centered in the sense that the human language comprehension system endeavours to identify the actor participant as quickly and unambiguously as possible while comprehending a sentence. Accordingly, if several candidates are available, they compete for the actor role and actor competition has measurable neurophysiological repercussions (Bornkessel-Schlesewsky and Schlesewsky 2009; Alday et al. 2014).

Actor identification in language processing is based both on domain-general features (e.g. animacy, certain movement parameters such as autonomous and/or biological motion, similarity to the first person etc.) and on language-specific features such as case marking or word order. In accordance with language-external observations regarding the importance of actor entities for mechanisms such as attentional orienting (New et al. 2007) or social cognition (Frith and Frith 2010), the eADM assumes that the actor can be viewed as a cognitive and neurobiological attractor category, with domain-general actor features allowing for the *bootstrapping* of language-specific actor characteristics during language development (Bornkessel-Schlesewsky and Schlesewsky 2014). Clearly, individual actor-related features will be more important for actor identification in certain languages as opposed to others (e.g. case marking in German, Japanese or Hindi versus English) and, within a particular language, some actor-related features will be weighted more strongly than others (Bates et al. 1982, 2001; Bates and MacWhinney, 1989; MacWhinney et al. 1984).

In this regard, parameter estimation – i.e. estimating the weighting of individual actor-related prominence features in a given language – becomes a central modelling problem. In the following, we introduce an initial, empirically-based framework for parameter estimation that is flexible, based on open source software components and thus freely distributable and requires only a minimal time commitment from test

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

subjects (i.e. native speakers of a given language). We thereby intend to establish a basis for examining (a) inter-individual differences in parameter weightings, and (b) lesser-studied languages for which only a small number of speakers is available to participate in linguistic experiments.

2.1 Previous computational work

Alday et al. (2014) presented the first computational implementation of actor competition, with a strong focus on distinctness (similarity/distance in the space of prominence features) as a predictor of mean EEG signal in time windows previously associated with actor competition. Due to high variance in EEG data – both inter- and intra subject – mixed effect models with crossed random factors for subjects and items were used. Moreover, the dependent variable was not a single offline behavioral measurement but rather an online measure of brain activity. The independent variables were different notions of distance, i.e. different mathematical ways of combining prominence features and weights into a single distinctness score. While models involving neurophysiological data are arguable much closer to the actual biological reality of language processing, they measure processes at a level where the correspondence between conscious intuition and subconscious computation is far from clear. As such, while the parameter estimation used here is of utmost importance for continued work on such models, the results of the two approaches are not directly comparable but rather complementary.

3 Individual experimentation

Robust parameter estimation must apply at the single subject (i.e. individual native speaker) level for several reasons. Language comprehension in a given language arguably involves a "strategy space" rather than hard-and-fast, deterministic processing strategy (see Howes, Lewis, and Vera (2009), for a more general cognitive perspective). Thus, by estimating inter-individual variability, we can establish an estimate of the breadth of the strategy space. Secondly, under certain circumstances (e.g. languages with few remaining or available speakers, clinical populations, children) it may not be possible to obtain data from a large pool of participants. Hence, the framework described here aims to provide a first step towards parameter estimation for individual participants, using the actor competition/prominence feature approach of the eADM as a test case. Of course, the approach is in principle applicable to any type of linguistic feature/model parameter.

3.1 Experiment

In order to maximize the portability and availability of individual parameter estimation, the experiment is restricted in equipment and duration. The experiment is programmed in OpenSesame (Mathôt et al. 2012), a freely available, Open Source software package written in Python for cognitive science experiments that runs on Windows, Mac OS X and Linux. No further equipment is required for the experiment itself. Similarly, the other parts of the proposed toolchain (R, Python and various packages for them) are all free software and available on all three platforms.

The much harder restriction is the duration of the experiment. While many psycholinguistic and neurolinguistic experiments last several hours per test subject, we restricted ourselves to a run time of between 30 and 40 minutes. This clearly restricts the number of trials available, which forces a tradeoff between a fully factorial exploration of differing conditions and the number of trials per condition. In the provided example experiment, the stimulus preparation script `load_data.py` generates the fullest factorial design allowed by its inputs (for our current sample stimuli, [ANIMACY × CASE × NUMBER] × [NP1, NP2], a total of 16 conditions, including violations) across all items and takes a random sample to generate 200 trials (see Table 1).

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

Table 1 Sample stimuli

dass	die Pfarrer	die Magier	angerempelt	haben.
that	*the pastors*	*the magicians*	*bumped-into*	*have.*
*dass	den Wirt	den Einbrecher	eingeladen	hat.
**that*	*the host.ACC*	*the thief.ACC*	*invited*	*has.*
dass	die Bürostühle	der Kellner	gespendet	hat.
that	*the office chairs*	*the waiter.NOM*	*donated*	*has.*
dass	die Zäune	die Tischler	bedauert	haben.
that	*the fences*	*the carpenters*	*regretted*	*have.*
dass	die Räume	der Veranstalter	eingeschaltet	hat.
that	*the rooms*	*the organizer.NOM*	*turned-on.*	*has.*
dass	den Bauer	der Obdachlose	gerettet	hat.
that	*the farmer.ACC*	*the homeless person.NOM*	*saved*	*has.*

Notes: All sentences began with *Gestern wurde erzählt* ("Yesterday, it was told"). Due to case syncretism in the German plural, all plural nouns were ambiguous and thus encoded in the subsequent models as being the average of nominative and accusative. The active task for the first sentence is _ *hat/haben angeremeplt* ("_s has/have bumped into"), with the two nouns placed on either side (left-right placement was random). Because the article in German carries most case information and some number information, it was omitted in the task.

This leads to an extremely sparse sample which will differ from run to run. The variation is in and of itself interesting, as it gives some indication of minimal learnability requirements.

The task for the experiment is a comprehension question, asking either for the actor or the undergoer ("Who did X?"/"Someone did X to whom?" or passive variants of the same).[1] For the syntactically ambiguous or ungrammatical sentences, this task attempts to force the subject to arrive at some interpretation of the sentence (as would happen in normal conversation). This serves as an explicit task somewhat similar to traditional acceptability judgements. The task is also timed with a moderately hard timeout of 4 seconds, which should push the subject to answer more intuitively and less metalinguistically. The answer is encoded as having assigned actorhood to the first or the second NP (cf. Bates et al. 1982; MacWhinney et al. 1984; Li et al. 1993; Kempe and MacWhinney 1999). "Correctness" is not a valid measure across conditions because the ambiguous and ungrammatical conditions lack a canonical answer. Moreover, the interesting question is how the prominence heuristic allows for decision under uncertainty. The response time is also recorded, under the assumption that prominence features misaligned with their weightings and the actor prototype will lead to higher reaction times.

3.2 Parameter estimation

Ultimately, the computational problem presented by the actor strategy is classification. The language system must assign actorhood to a single argument and, in order to do that, depends on classifying an individual argument's probability of being an actor. Probabilitistic classification into two groups is a well-researched problem with many methods available. The simplest method, based on the general linear model, is probit regression (Bliss 1934).

The dependent variable in probit regression is a probabilisitic binary classification, while the independent variables are the feature encodings.[2] The model weights correspond to the coefficient estimates, allowing direct extraction of the weights and easy interpretation. Although the better known logistic

1 In initial tests, it seems that test-subjects felt more comfortable when only active or passive questions were presented. The results for one volunteer who completed both the mixed and pure passive variants were similar, resulting in the same rankings. However, the two trial runs are not directly comparable because each run used a different subset of the possible stimuli.

2 Currently, the features are encoded as binary pairs with 1 for marked/more prominent and 0 for less prominent. For many things, a discrete scale seems unnatural, and the model can accommodate continuous scales on [0,1] without modification. This is currently used for ambiguous case marking, encoded as 0.5.

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

regression yields similar results and has coefficients that are slightly easier to interpret (as an odds ratio), probit regression has several advantages for modelling the role of prominence features.[3] Logistic regression is more difficult to implement in a Bayesian framework, which is a disadvantage for planned future work utilizing estimates from several sources (i.e. work involving non-flat priors and pooling). More importantly, probit regression is the better model for the dichotomization of a continuous variable. Prominence is a continuous variable, but the actor strategy is a dichotomization strategy,[4] and probit regression is thus better suited.[5]

For a transitive relationship, there are two arguments each carrying their own set of prominence features. While it is possible that the weights for the features are position dependent (i.e. that there is an interaction term for argument position by prominence), we make the simplifying assumption that this is not the case. Accordingly, we can collapse the two sets of features into a single set of pairwise differences, thus reducing the number of parameters to be estimated. This also makes the work more compatible with the types of models used in Alday et al. (2014), where the weights actually applied to the pairwise differences. Here we use NP1 − NP2 to model expectations: NP1 − NP2 < 0 means that a more prominent argument comes in a later position, which is known to be dispreferred (preference for initial actors).

Due to the sparseness of the data, we also exclude interaction terms between features.

3.3 Sample analysis

3.3.1 Actor identification

Sample analysis scripts and data sets collected from graduate students are included in the supplementary materials. In the following, we present the results from `sample01a`.[6]

For our regression, a 1 encodes an initial actor, while a 0 encodes an initial undergoer. Because of the mutual exclusivity of the actor-undergoer relation, we do not need to encode the other argument. We chose one to correspond to initial actor so that more prominence would correlate positively with more actorhood. Table 2 provides the results of the regression with `sample01a`.

Table 2 **AIC**: 215.6 **Deviance**: 205.6 **Null Deviance**: 263.7 Results of probit regression for actor-initial order

	Estimate	Std. Error	z value	Pr(> |z|)
(Intercept)	0.2878	0.2064	1.4	0.163
animacy	0.6269	0.167	3.8	0.0001743
case	1.3480	0.2318	5.8	6.014e-09
number	−0.4678	0.1596	−2.9	0.003381
index	0.0005	0.001767	0.3	0.7614

Notes: The large residual deviance reflects the great variety of items, including syntactic und semantic violation conditions

3 The model coefficients in logit regression correspond to changes in the odds ratio for the dependent variable per unit change in the independent variable. In probit regression, the coefficients represent change in the *z*-score of the dependent variable per unit change in the independent variable. The errors in logistic regression follow the logistic function, while the errors in probit regression follow the normal distribution. This leads to the "tipping" behavior of emergent binary categories from a continuous scale. Both methods depend on a logarithmic link function.

4 Indeed, the currently postulated processing model assumes that the threshold for this dichotomization is dynamic, adapting to contextual demands.

5 To our knowledge, no psycholinguistic study has utilized probit regression. However, logistic regression, which is generally better known, and its mixed-effect extension have been proposed as a more appropriate way to analyze categorical responses, cf. Jaeger (2008) and others.

6 The number refers to the test subject, while the a refers to the task (active question, i.e. name the actor). Other possible codes are p (passive question, i.e. name the undergoer) and b (both types of questions randomly mixed).

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

While the exact meaning of the estimates in probit regression is difficult, the relationship in the size of the estimates is straight forward. Case clearly has the strongest estimate, which fits well given that unambiguous case marking is known to work deterministically in German. Case also has the least amount of variance relative to its influence – this is reflected in the large *z*-score. Animacy has the next highest estimate and *z*-score, with both about half as large as for case. Number has the estimate and *z*-score with the smallest magnitude. Interestingly, the sign is also reversed for number. This could reflect the late disambiguating nature of number agreement in the verb final sentences. Index (i.e. trial number within the experiment) has a very small estimate and *z*-score, which indicates that the test subject was unable to develop and apply a strategy during the course of the experiment.[7]

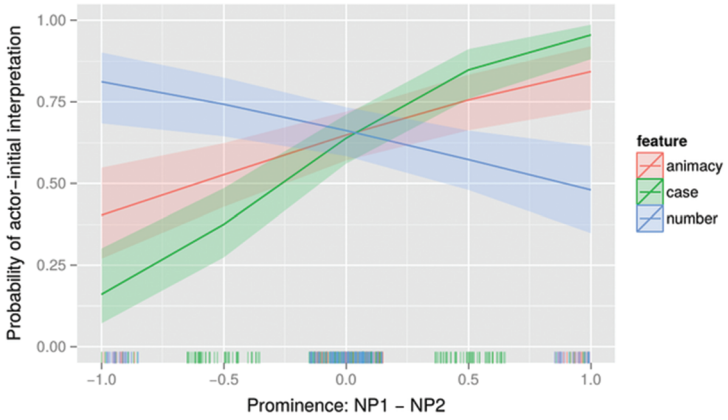

Figure 1 Actor identification by feature

This is also clear graphically. Figure 1 presents the likelihood of choosing an initial actor based on the difference in prominence. High initial prominence followed by low initial prominence – the high end of the scale – increases the odds of assigning actorhood to the first argument. (Shaded regions indicate the 95% confidence interval; the number of samples of each condition is shown as a rug plot.) The strength of a cue is reflected in the slope of the individual lines. The preference for an initial actor can also be clearly seen here. At 0, i.e. at a tie in prominence between the two arguments, all features show a preference for an actor-initial interpretation. Despite the low power from a short, unbalanced design, a clear ranking is visible.

The emergence of such a clear ranking is also interesting for exploring the learnability of the actor strategy. Figure 2 shows the convergence of the parameter estimates by recomputing the model for the first *n* trials, starting with trial 25. Despite the low statistical power in such an experiment, the parameter estimates quickly sort themselves into a clear ranking. Figure 3 shows how this learning would look in terms of language processing, displaying the identification curve as in Figure 1, but after 50, 100, 150 and 200 trials. Again, after 50 trials, the strength of case is established, but the strength of the next strongest cue, animacy, is established after twice as many (100), and the third strongest cue settles down between 150 and 200 trials (1.5–2x as many as animacy), suggesting perhaps a rank-power law in cue strength (such as in Zipf's Laws).

7 This of course says nothing about whether the test subject developed one beforehand while looking at an example run.

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

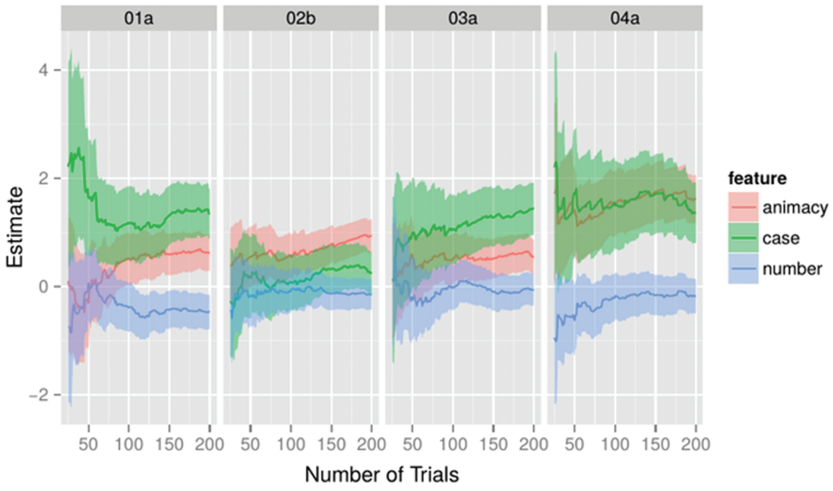

Figure 2 Convergence of estimates

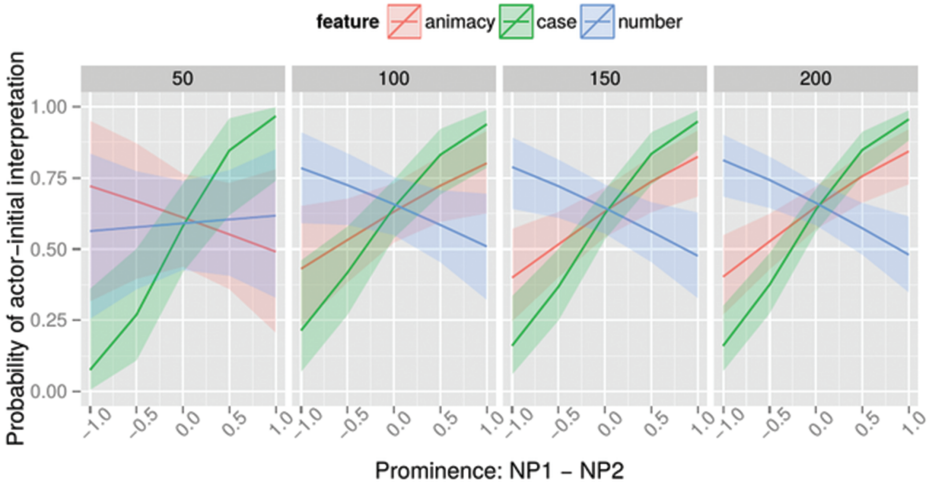

Figure 3 Effect of estimate precision on actor identification by feature

Although we did not model separate interaction terms here, it is nonetheless interesting to consider how the model handles interaction. For this, we used the model to predict outputs for the grid of animacy × case, treating both as semi-continuous measures. Figure 4 shows the contour for this simulation holding constant index = 1 and number = 0 (plural). Darker tones indicate a higher probability of initial actor interpretation.[8] Using a physical metaphor, we can view the darker tones as being valleys and the lighter tones as being hills. Actorhood works as an attractor, with the ideal attractor basin being an initial, animate, nominative argument. However, even an inanimate initial nominative is more likely to be interpreted as an actor than an animate initial accusative – the basin slopes more sharply along case than along animacy.

8 Because the calculated model includes index and number as independent variables, they have to be nominally set. Index has little effect and so the choice is completely arbitrary: 1 is reasonable because it reflects a neutral state (not influenced by the odd experimental context) and is the smallest valuable possible. Similarly, the effect for number was quite small and the choice is again arbitrary. Number = 0 (plural) reflects the base state.

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

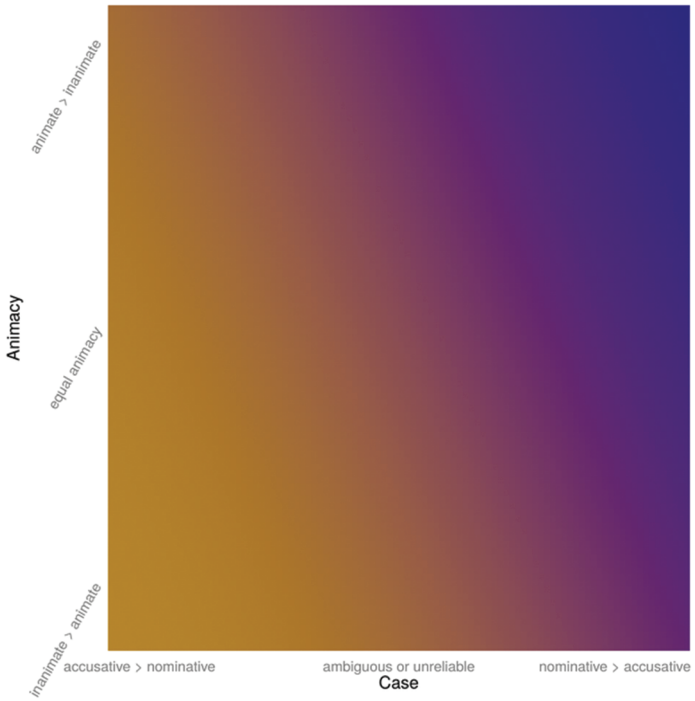

Figure 4 Individual actor space

Figures and models for four test subjects can be found in the Appendix.

3.3.2 Reaction time

The reaction time model (Table 3) shows the same general trend in the magnitude of the various estimates[9] but much higher variance. This variance leads to poor *t*-values. Index has a very small effect here – roughly 2 ms reduction in reaction time for each successive item. This reflects a "training effect", where the test subject adapts to the experimental conditions and task. Although this effect is *statistically* highly significant, the magnitude is quite small.

Table 3 **AIC**: 2973.8 **Adjusted R^2**: 0.1 **Residual standard error**: 402.7 on 195 degrees of freedom, $F(4, 195) = 3.96$, $p = 0.0041$. Results of linear regression model for reaction time

	Estimate	Std. Error	*t* value	Pr(> |*t*|)
(Intercept)	863.2570	57.57	15	2.618e-34
animacy	−48.1050	44.68	−1.1	0.283
case	−66.2133	56.59	−1.2	0.2434
number	−16.8376	42.74	−0.39	0.694
index	−1.8386	0.4956	−3.7	0.0002706

The high degree of variance in the reaction time (related to the complexity of the task) and the limited power of such a small experiment leads to promising yet not reliable results. Training has a significant effect on test subject ability: a roughly 400 ms reduction over the course of the experiment. Over a larger

9 The reversed sign reflects that increased prominence aligns better with general expectations (actor preference) and thus *reduced* reaction times.

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

experiment, we expect that this effect would reach some natural asymptote as the test subjects become comfortable with the task and that the accompanying reduction in the variance would lead to larger *t*-values.

More research is required in order to investigate how best to integrate reaction time into the parameter estimation.

3.4 Comparison to previous work

Using the model estimates collected here, it is possible to compare empirical weights with the a priori estimates presented in Alday et al. (2014). More precisely, the weight of a given feature is given by e^{β}, where β is the coefficient in the probit model. The exponentiation is necessary because the probit link function is logarithmic. We can also compute a mixed effects model over the four subjects tested thus far and extract the fixed-effect relevant coefficient for the pooled weightings. A comparison of the **apriori**, single-subject **emp**irical and **pooled** models of the best distinctness measure (sdiff, Alday et al. 2014) is presented in Table 4. Critically, although two different sets of test subjects were used, the models are all extremely similar in their fit. The *t*-score for the distinctness metric was also similar.

Table 4 Comparison of sdiff performance for a priori weights, weights from `sample01a`, and all samples. Models fitted to the EEG data in the N400 window (Alday et al. 2014)

	df	AIC	BIC	logLik
model.apriori	11	395190	395291	−197584
model.emp	11	395223	395324	−197600
model.pooled	11	395252	395352	−197615

4 Better data through openness

An exciting aspect of estimating model parameters based on individual performance in a quick experiment is the possibility of making science accessible, available and touchable to everyone, which should open the door for exploration of areas where data acquisition has been difficult, such as the study of individual differences and less-researched languages. This depends on the software and underlying methodology being freely accessible, free to modify and free to distribute. All software used here is licensed under the GNU Public License (GPL). For the portions we wrote, we encourage you to fork us on Bitbucket: bug fixes and improvements are of course welcome, but example stimuli for different languages, sample data and alternative analyses would contribute far more towards our and the broader community's understanding of language.

4.1 Future plans

Our own future plans for the software include a more integrated tool chain. Currently, the user has to install several programs (Python + several extensions, OpenSesame, R), but it should be possible to move core features into the OpenSesame experiment. The user would perhaps no longer have access to more advanced features (for which she would need some programming know-how anyway), but a core set of features for spontaneously testing a single-subject would fit into a single OpenSesame experiment file pool. As part of

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

this, we are currently implementing a framework for combining the estimated parameters with an existing computational framework and providing an animation of how sentence processing in an individual works. All models are wrong, including ours, but some are useful (Box and Draper 1987) and the most useful are the ones everybody can see and tinker with.

4.2 More data, less uncertainty

The framework presented here shows that parameter estimation is possible even with few trials from a single subject. With minimal equipment and quick parameter estimation, it is now possible to gather data from more languages, and we have another tool to remove our Indo-European blinders. At the other end of the spectrum, model fit may be poorer with a few test subjects than with a single subject, if the variance between subjects is large. This would be an interesting result within itself, indicating the size of the strategy space for a single language (population). More data from more languages and more individuals will help us to better understand both the cognitive mechanisms underlying language in general and the speaker-level adaption to a particular language.

Acknowledgements: Parts of the research reported here were supported by the LOEWE programme funded by the German state of Hesse. The authors would like to thank Alexander Dröge for contributing the stimuli for the German experiment as well as Jona Sassenhagen and Elisabeth Rabs for their extensive help in testing the experiment. Jona Sassenhagen also gave extensive feedback on the data visualization. Miriam Burk, Christina Lubinus, Pia Schoknecht and Fiona Weiß kindly consented to us using their performance as sample data.

Appendix

In the following, we provide additional figures comparing four subjects.

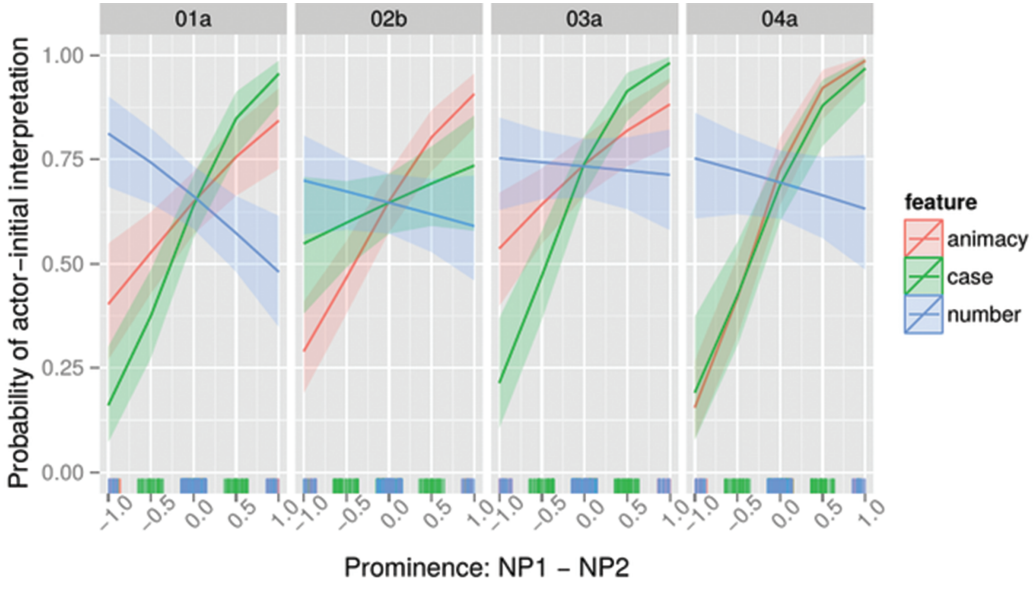

Figure 5 Actor identification by feature

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

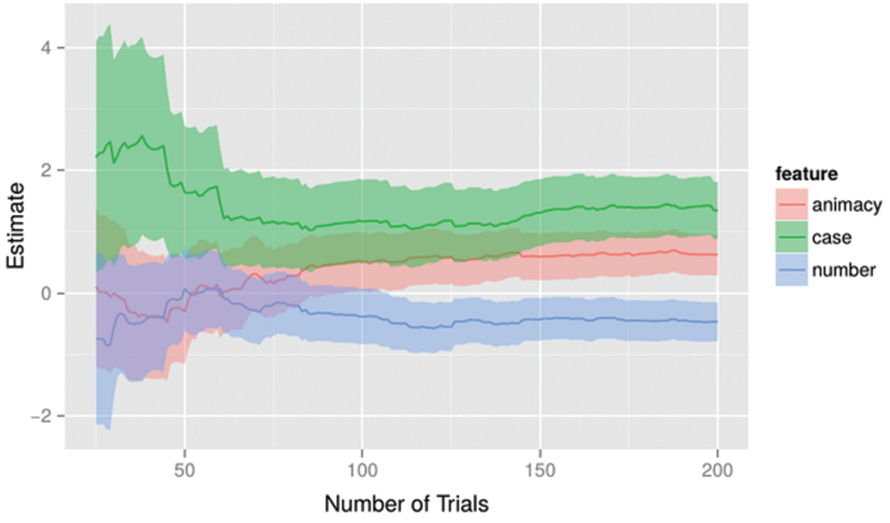

Figure 6 Convergence of estimates

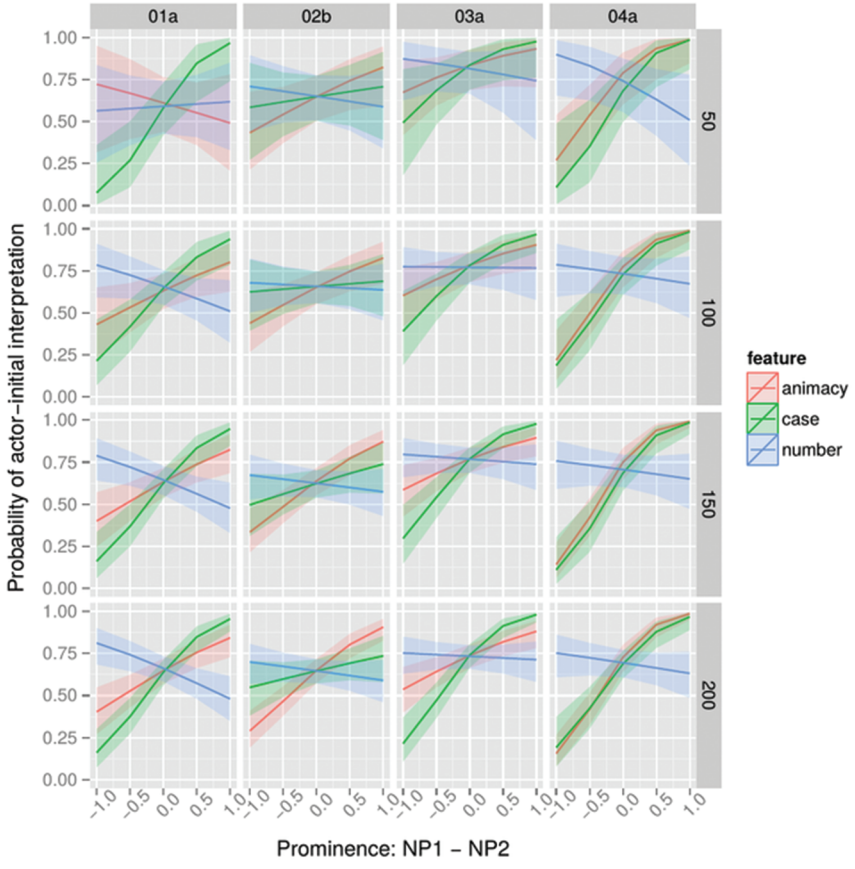

Figure 7 Effect of estimate precision on actor identification by feature

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

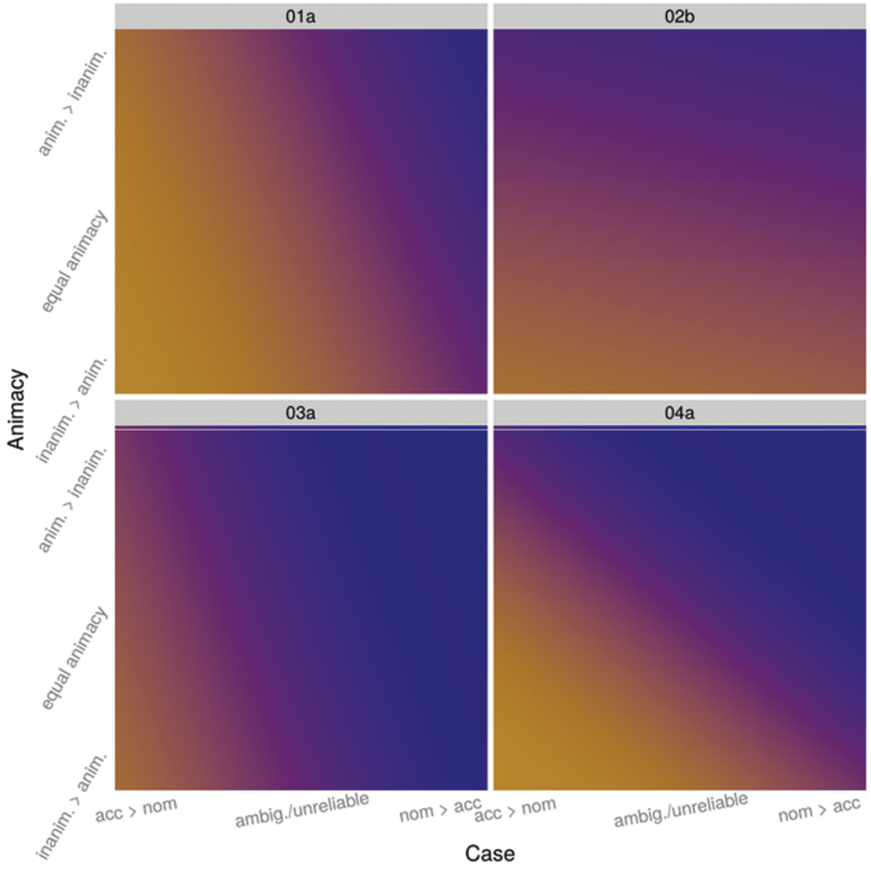

Figure 8 Individual actor space

References

Alday, Phillip M., Matthias Schlesewsky & Ina Bornkessel-Schlesewsky. 2014. Towards a computational model of actor-based language comprehension. *Neuroinformatics* 12(1). 143–179.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2014. *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7*. https://github.com/lme4/lme4/

Bates, Elizabeth, Antonella Devescovi & Beverly Wulfeck. 2001. Psycholinguistics: A cross-language perspective. *Annual Review of Psychology* 52. 369–396.

Bates, Elizabeth & Brian MacWhinney. 1989. Cross-linguistic research in language acquisition and language processing. In *Proceedings of the World Conference on Basque Language and Culture*. San Sebastian: Basque Regional Government.

Bates, Elizabeth, Sandra McNew, Brian MacWhinney, Antonella Devescovi & Stan Smith. 1982. Functional constraints on sentence processing a cross-linguistic study. *Cognition* 11. 245–299.

Bliss, C.I. 1934. The methods of probits. *Science* 79(2037). 38–39. doi:10.1126/science.79.2037.38.

Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky. 2006. The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review* 113(4). 787–821.

Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky. 2009. The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass* 3(1). 19–58.

Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky. 2013. Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension." *Brain and Language* 125. 60–76.

Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky. 2014. Competition in argument interpretation: Evidence from the neurobiology of language. In Brian MacWhinney, Andrej Malchukov & Edith Moravcsik (eds.), *Competing motivations in grammar and usage*, 107–126. Oxford: Oxford University Press.

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM

Box, George E.P. & Norman R. Draper. 1987. *Empirical model-building and response surfaces*. Probability and Mathematical Statistics. Oxford: John Wiley Sons.

Dahl, David B. 2014. *xtable: Export tables to LaTeX or HTML*. R package version 1.7-4. http://CRAN.R-project.org/package=xtable

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619. Linguistic Society of America.

Fox, John. 2003. Effect displays in R for generalised linear models. *Journal of Statistical Software* 8(15). 1–27.

Fox, John & Jangman Hong. 2009. Effect displays in R for multinomial and proportional-odds logit models: Extensions to the effects package. *Journal of Statistical Software* 32(1). 1–24.

Frith, U & C.D. Frith. 2010. "The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society B* 365. 165–176.

Howes, Andrew, Richard L Lewis & Alonso Vera. 2009. Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review* 116(4). 717–751. doi:10.1037/a0017187.

Jaeger, T Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59(4). 434–446. doi:10.1016/j.jml.2007.11.007.

Kempe, Vera & Brian MacWhinney. 1999. Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes* 14(2). 129–171.

Li, Ping, Elizabeth Bates & Brian MacWhinney. 1993. Processing a language without inflections: A reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language(Print)* 32(2). 169–192.

MacWhinney, Brian, Elizabeth Bates & Reinhold Kliegl. 1984. Cue validity and sentence interpretation in English, German and Italian. *Journal of Verbal Learning and Verbal Behavior* 23(2). 127–150.

Mathôt, Sebastiaan, Daniel Schreij & Jan Theeuwes. 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* 44(2). 1–11.

New, Joshua, Leda Cosmides & John Tooby. 2007. Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences* 104(42). 16598–16603. doi:10.1073/pnas.0703913104.

Primus, Beatrice. 1999. *Cases and thematic roles*. Tübingen: Niemeyer.

Simmons, Joseph P., Leif D. Nelson & Uri Simonsohn. 2011. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science* 22(11). 1359–1366.

Team, R Core. 2014. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Van Valin, Robert D. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

Wickham, Hadley. 2007. Reshaping data with the reshape package. *Journal of Statistical Software* 21(12). 1–20.

Wickham, Hadley. 2009. *ggplot2. Use R!* New York: Springer.

Brought to you by | MPI fuer Psycholinguistik
Authenticated
Download Date | 7/25/17 5:00 PM