# Robots That Say 'No'

Frank Förster, Chrystopher L. Nehaniv, and Joe Saunders

Adaptive Systems Research Group, School of Computer Science
University of Hertfordshire,College Lane, Hatfield, AL10 9AB.United Kingdom

**Abstract.** This paper* reports on foundational considerations for experiments into the acquisition of human-like use and understanding of negation in linguistic utterances via a developmental robotics approach. For this purpose different taxonomies of negation in early child language are analysed in order to show the large variety of communicative functions that these different types of negation have. Requirements for robotic systems that aim at acquiring these utterances in a linguistically unconstrained human-robot dialog are derived from this analysis.

## 1 Introduction

This paper presents an analysis of negation in early child language which offers an alternative to the prevailing paradigm of propositional representation with one that considers language as a means to manipulate the world. Apart from being necessary in order to ground negative utterances this shift in the viewpoint seems more in accordance with evolutionary perspectives on language [9]. The discussion below contributes to ongoing research into achieving human-like language acquisition in robots via developmental processes [16, 10, 5].
The motivation for the development of existing frameworks is at least partially to tackle the symbol grounding problem by linking the symbolic representations of the system to sensorimotor data [8, 15]. Some of the embodied frameworks such as [19] enable artificial agents to invent their own vocabulary and simple forms of grammar like word order for the purpose of communicating with each other in a language constructed by and understandable to the robotic participants of the dialog [18]. Other frameworks, e.g. [6] or [14], enable robots to learn and understand names for objects, actions, or spatial relations, simple propositions, or commands taken from natural human language. From the perspective of pragmatics all of these frameworks enable robots to engage in one or two types of speech acts: commenting on the state of the world, an assertive speech act, and following orders, a reaction to directive speech acts. As is known from speech act theory [1, 17] human language has many more functions than only stating facts about the world or receiving and giving orders. Observations of the earliest language use of children show clearly that the *functional* scope of

* **Acknowledgement** The work described in this paper was conducted within the EU Integrated Project ITALK ("Integration and Transfer of Action and Language in Robots") funded by the European Commission under contract number FP-7-214668.

early human language transcends these types of speech acts [4]. Already in the pre-grammatical phase of early utterances children express linguistically attraction or aversion towards objects, actions and persons by requesting or refusing. The latter constitutes only an alternative means for conveying communicative intentions which in the pre-linguistic phase were conveyed via gestures or prosodically marked non-lexical utterances that accomplish the same function. Negatives play a major role in so called "langage de volonté" (language of will) [7] and many types of negation are prototypical examples for utterances that cannot be grounded without being linked to the volitional or affective state of an agent. Pea [13] uses the apt expression of "motor-affective sensorimotor intelligence" to characterize the capability that children must have in order to refuse things in a linguistic manner. This highlights the difference compared to pure sensorimotor grounding of object words like "cup" or "ball" or action words like "grasp" or "throw". Prohibitive speech acts, acts of refusal, or acts of motivation-dependent denial are examples for negatives that are at least partially distinguished through their affective nature. Other negative speech acts include commenting on disappearance of objects, cessation of events, or expressing unfulfilled expectations which could be grounded with the existing frameworks. It can neither be assumed that the last mentioned types constitute the overwhelming majority of early negation nor can these be expected to be the dominant types of negation in a human-robot dialog with unconstrained language use on part of the human.

We examine the preconditions of this very early form of human speech that must be met by robotic agents in order to ground negatives and engage in negative speech acts in the context of such a dialog. Less affective utterances that comment on disappearance or express unfulfilled expectations, sometimes grouped together under the term of non-existence, have the advantage of being a natural counterpart to positive comments on and descriptions of the world and thus highlight the difference between recent frameworks and the new approach proposed here, which aims at extending these frameworks towards the grounding of negation including affect-heavy forms like rejection or prohibition.

The presented outline is meant to be a theoretical basis for research which is dedicated to enable robots to obtain early pre-grammatical human language. The account tries to stay close to early human language development while at the same time exposing the crucial properties of negation which have to be met by robotic systems in order to be able to engage in dialogues with humans. Section 2 gives a short overview of negative speech acts observed in early child language in the one-word stage and their characteristics. Different taxonomies are presented and compared with specific regard to properties that are crucial for grounding and representation of negatives in robotic language acquisition. These acts will be considered from the point of view of pragmatics, speech act theory, semiotics, and accounts of early language acquisition [4, 12]. Section 3 outlines implications of the linguistic analysis for robotic frameworks and describes cornerstones of a computational model designed to enable humanoid robots to acquire early human language via dialog with a human partner. Section 4 concludes with a short summary.

## 2 Emergence of negation in human language development

Compared to words that name objects, actions or even abstract categories, negation is semantically hard to classify and is deeply interwoven with lexical meaning in general [13]. There are several strategies to categorize negation semantically with varying granularity (see also [13] for a discussion). To the knowledge of the authors there is despite many centuries of research on many aspects of negation no commonly accepted taxonomy with regard to its semantics. This circumstance indicates the difficulty of this enterprise. We adopt in the following the taxonomy proposed by Pea [13] as it is motivated by similarities in the situational context and the child's behavior which leads to an intuitive clustering. Moreover this taxonomy exhibits the nice feature of emphasizing different properties for the particular negation types that lend themselves readily to a translation into different requirements for the computational model delineated in section 3 below. Note that this taxonomy does not necessarily reflect the child's own distinction (see [13]). Other taxonomies can be found in [2] and [3] and will be compared to the adopted one. Table 1 summarizes the most frequent types of early negation and their characteristics in the period up to 25 months according to Pea [13]. The negation types in the table roughly emerge in the order listed, whereby the three middle types can also be exchanged with regard to order of emergence. The reason for this variability is simply that different parents from potentially different social classes speak very differently to their children which has a strong impact on their linguistic development. Nonetheless the tendency from a rather reactive behavior towards a behavior that requires an increasingly complex memory is observable.

**Negation types not listed in table 1** Less frequent and therefore omitted types are make-believe, agreement to negative statements, motivation-dependent, and perspective-dependent denial (see [13]).

**Rejection** Action-based negations that serve to reject objects, persons, activities or events in the immediate environment fall into this category. Often expressed as a simple "No" this type of negation is repeatedly reported to be the first type that emerges and has gestural and non-gestural predecessors long before the emergence of the first word. Rejections might be the most affective type of negation as they cannot be interpreted without reference to aversion.

**Self-Prohibition** Utterances of negation in the context of objects or actions that have been forbidden previously by the teacher. One can observe these utterances in a scenario, where the child approaches a forbidden object, hesitates to touch it while looking at the teacher, approaches it again and so on. Self-prohibition is more complex than rejection in the sense that it requires a internal representation of the preceding external prohibition.

**Disappearance** Typically expressed as "gone" or "all-gone" in English, disappearance negatives signal the disappearance of something in the immediate past. Like self-prohibition this type requires an internal representation, in this case of objects that disappear, but on a shorter time-scale. Short-term memory might

| Negation type | Topic | Characteristics |
|---|---|---|
| Rejection | objects, persons, events, activities in the present | affective, adjacent and nonadjacent, action-based, no need for internal representation |
| Self-Prohibition | objects, persons, events, activities | nonadjacent, (affective)*, referent in present but assumes prohibition in past with regard to the same referent on part of the carer |
| Disappearance | objects, persons, events, activities | adjacent and nonadjacent, non-affective*, referent in the immediate past, typically in context of where-questions but also with declaratives, need for internal representation |
| Unfulfilled Expectation | objects, persons, events, activities | nonadjacent, (affective)*, referent in past or present, comment on constraint on activity or absence other than immediately prior disappearance or cessation, need for internal representation, first action-uses then existential and locational uses |
| Truth-functional Denial | facts of the situation which the proposition that is to be denied refers to | adjacent, non-affective*, response to a proposition, need for abstract internal representation |

Table 1: Most frequent types of early negation. Taxonomy and characteristics are compiled from Pea [13] except for characteristics tagged with * which were added by authors. (affective) means that corresponding negation types are not intrinsically affective but accompanied or caused by affective states. 'Adjacent' describes whether or not an utterance is produced in response to another one.

be enough to support this type.

**Unfulfilled Expectation** Like disappearance, unfulfilled expectations might be expressed as "gone" or "all-gone". They are uttered in contexts where objects are absent from their expected or habitual location without having been present in the immediate past. They also occur when an activity is unsuccessful in contrast to previous success, e.g. caused by broken toys. This type requires representations of objects, actions or events on longer time-scales than disappearance.

**Truth-functional Denial** This type of negation is the most abstract in the taxonomy and the last to emerge. Generally it is a response to a proposition not held to be true by the child. To employ this kind of negation a child must be able to conduct logical judgments and use at least some truth-conditional semantics of language. The negated proposition is independent of the child's attitude towards it, distinguishing this from motivation- and perspective-dependent denial. The facts that are referred to might be in the present, past or future.

**Taxonomy proposed by Choi** Choi divides cross-linguistic negation up to age 40 months into nine categories, which roughly emerge in three developmental phases [3]:
**Phase 1**: (nonexistence), prohibition, rejection, (failure)
**Phase 2**: denial, (inability, epistemic negation)
**Phase 3**: normative negation, inferential negation
The brackets indicate a variation in the time of emergence for the different children observed by Choi. For some children they emerged during the indicated phase for others they emerged one phase later. With *normative negation* Choi describes negatives that occur when the state of affairs differs from the agents habitual expectations. Negations of this type are evoked through a deviation from normative expectation (e.g. persons go *in* and not *on* a car) or the unorthodox use of a tool. *Inferential negation* is related to denial, but in contrast the agent assumes that the conversation partner holds the statement which is to be denied to be true rather than having actually heard the latter expressing it. *Nonexistence* is expressed when the agent expects an entity to be present which is not or the entity disappears. Nonexistence according to Choi therefore seems to subsume Pea's categories of disappearance and unfulfilled expectation. *Prohibition* is used by the agent to negate action on part of the interaction partner. Pea does not list this category explicitly but notes that it is tightly linked to rejection. *Failure* is the reaction to a specific event that does not occur as expected. Like nonexistence it therefore seems to map to Pea's category of unfulfilled expectation. *Denial* probably subsumes all the differentiated types of denial of Pea. On the other hand Choi only quotes one example which falls into Pea's category of truth-functional denial. *Inability* describes an agent's negation of its physical ability to accomplish a task. It probably maps best to Pea's unfulfilled expectation category as this also subsumes constraints on activities. *Epistemic negation* describes negative responses to requests for information like "I don't know". This type does not seem to be captured by Pea's taxonomy maybe because he did not observe it while monitoring the children of his study.

**Taxonomy proposed by Bloom** Bloom is cited by both of the other authors and seems to have provided one of the first accounts of the development of negation with regard to semantics [2]. At the same time her partition is the least elaborate one of the presented taxonomies. She only distinguishes the three types nonexistence, rejection and denial. Another striking difference is the fact that she reports nonexistence as the first negation type to emerge. Pea argues that this might be the case because Bloom focuses on negative meanings in sentences instead of their emergence during the single-word period.
**Nonexistence**: The referent is not manifest in the context, where there is an expectation of its existence, and is correspondingly negated in the linguistic expression.
**Rejection**: The referent actually exists or is imminent within the contextual space of the speech event and is rejected or opposed by the child.
**Denial**: The negative utterance asserts that an actual (or supposed) predication is not the case. The negated referent is not actually manifest in the context as

it is in *rejection*, but it is manifest symbolically in a previous utterance.

Bloom's category of *denial* is evidently the same as Pea's category of truth-functional denial. *Rejection* is defined in the same way as given by Pea. *Nonexistence* maps to Pea's unfulfilled expectation. Blooms focus on sentential negation renders it incomparable with regard to developmental emergence and we therefore will not take it into account in what follows.

## 3 Prerequisites for grumpy robots

From the linguistic analysis above we infer that there are three crucial properties that distinguish the different types of negation. First they are distinguished through their relatedness to affect or volition. E.g. rejection cannot be grounded meaningfully without referring to the affective state of the agent. We therefore propose to replace sensorimotor grounding with sensorimotor-affective grounding to take this circumstance into account. An easy way to accomplish this might be to simply introduce two-dimensional values for affect denoting the valence (positive/negative) and the degree of affect. These values principally could be associated to instances of acquired concepts, linked to ongoing 'needs' of the robot, treated in the same way as other sensorimotor values and stored together with the other sensor readings for each frame in two additional dimensions. These values constituting the willingness to cooperate/accept or the opposite thereof could in the beginning be chosen arbitrarily indicating a tendency to accept or reject certain actions and objects. For the sole purpose of grounding the reason why an affective state is the way it is seems not of importance. In order to signal the state of affect to the interaction partner in language acquisition games this state should be mirrored by facial or body gestures of the humanoid. This is necessary in order to provoke negative utterances by the interaction partner in the case of non-cooperation/non-acceptance. This could be utterances like "No? You don't like this ball?". In later stages the arbitrary choice of affect could be replaced by more sensible measures like the outputs of a planning module depending on the decision if a certain action is useful or harmful to the agent to achieve its goals, to satisfy its internal drives, or to maintain its state [11].

The second distinction we can observe is the increasing complexity with regard to the required memory: from a purely reactive behaviour in the case of rejection to the need for long-term memory for unfulfilled expectations or normative negation. Interesting issues in this context are the need for an internalization of a previously external physical prohibition in the case of self-prohibition and the need for an internal representation of habitual locations and habitual functionality in the case of unfulfilled expectations. It is unclear whether these two memory-related requirements should be treated in a differentiated manner or whether it would be advantageous to distinguish them on a higher level but map them to the same underlying memory-structure. Experiments with different memory-models in which both negation types are acquired simultaneously should shed light onto this issue.

**Fig. 1.** Humanoids iCub (left) and Kaspar (right) used in language acquisition studies. These studies include work on learning vocabulary, holophrases, negation and grammar in a manner where language is grounded in active manipulation of objects, the social environment, and in sensorimotor experience in unconstrained interaction with human teachers. The work is targeted at the acquisition of human language-like capabilities.

The third distinction for the types is given by their property of being adjacent or nonadjacent (see table 1). In order to engage in adjacent negation the agent must know when it is addressed. The only purely adjacent type is truth-functional negation, the latest negation type expected to emerge. In this case turn-taking is an issue and the analysis of prosody or an artificial replacement thereof as the agent is supposed to react to being spoken to and therefore must have the means to find out when it is addressed.
The large functional variability of the different negation types suggests that from a developmental point of view early negation might be rather considered as a family of related but different types of speech acts than being a variation of a single phenomenon. Seen from this perspective an implementation in the spirit of item-based constructions seems natural [20]. The latter support a treatment of these types as if they were entirely independent in the beginning and leave it up to machine learning algorithms to detect the similarity. Eventually a higher-level schema constructed through these mechanisms might emerge that could be labeled negation by an external observer.

## 4 Summary

We provided an analysis of early negation types with regards to their functional use which highlighted crucial differences with regards to their treatment in the context of robotic language acquisition in general and furthermore in a setting of human-robot dialogues. We propose to introduce a representation of affect or volition into the system in order to enable the acquisition of particular negation types like rejection. We also propose to introduce a differentiated memory architecture in order to support other types of negation like disappearance or unfulfilled expectations. The fact that all but one negation type are expressed

also in a nonadjacent manner suggests that the integration of means to detect questions, e.g. through prosody extraction, is of importance but seems to be not of equal priority. Nonetheless turn-taking and the detection of questions is important in order to maintain the progress of the dialog which drives the grounded acquisition process itself. Initially it could be replaced by easier means of dialog control until e.g. robust methods for social cue and prosody detection are found.

## References


1. J. L. Austin. *How to Do Things with Words*. Harvard University Press, 1975.
2. L. Bloom. *Language Development: Form and Function in Emerging Grammars*. M.I.T. Press, 1970.
3. S. Choi. The semantic development of negation: a cross-linguistic longitudinal study. *Journal of Child Language*, 15(3):517–531, 1988.
4. E. V. Clark. *First Language Acquisition*. Cambridge University Press, 2009.
5. P. F. Dominey, G. Metta, F. Nori, and L. Natale. Anticipation and initiative in human-humanoid interaction. In *8th IEEE-RAS Intl. Conf. Humanoid Robots*, pages 693–699, 2008.
6. P.F. Dominey and J.D. Boucher. Learning to talk about events from narrated video in a construction grammar framework. *Artificial Intelligence*, 167:31–61, 2005.
7. P. Guillaume. Les débuts de la phrase dans le langage de l'enfant. *Journal de Psychologie Normale et Pathologique*, 24:1–25, 1927.
8. S. Harnad. The symbol grounding problem. *Physica D*, 42:335–346, 1990.
9. J. Maynard Smith and E. Szathmáry. *The Major Transitions in Evolution*. Oxford University Press, 2002.
10. Y. Nagai and K.J. Rohlfing. Computational analysis of motionese toward scaffolding robot action learning. *IEEE Trans. Autonomous Mental Development*, 1(1), 2009.
11. C. L. Nehaniv, K. Dautenhahn, and M. J. Loomes. Constructive biology and approaches to temporal grounding in post-reactive robotics. In G. T. McKee and P. Schenker, editors, *Sensor Fusion and Decentralized Control in Robotics Systems II*, volume 3839 of *Proc. of SPIE*, pages 156–167, 1999.
12. R. E. Owens. *Language Development: An Introduction*. Pearson Education, 2007.
13. R. Pea. The development of negation in early child language. In D.R. Olson, editor, *The Social Foundations of Language & Thought: Essays in Honor of Jerome S. Bruner*, pages 156–186. W.W. Norton, 1980.
14. D. Roy. Semiotic schemas: A framework for grounding language in action and perception. *Artificial Intelligence*, 167(1–2):170–205, 2005.
15. D. Roy and E. Reiter. Connecting language to the world. *Artificial Intelligence*, 167:1–12, 2005.
16. J. Saunders, C. Lyon, F. Förster, C.L. Nehaniv, and K. Dautenhahn. A constructivist approach to robot language learning via simulated babbling and holophrase extraction. In *Proc. 2nd Intl. IEEE Symp. on Artificial Life*, pages 13–20, 2009.
17. J. R. Searle. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, 1969.
18. L. Steels. The origins of syntax in visually grounded robotic agents. *Artificial Intelligence*, 103(1-2):133–156, 1998.
19. L. Steels. Evolving grounded communication for robots. *Trends in Cognitive Science*, 7(7):308–312, 2003.
20. M. Tomasello. *Constructing a Language*. Harvard University Press, 2003.