Recurrent Neural Network Sentence Parser for Multiple Languages with Flexible Meaning Representations for Home Scenarios

Xavier Hinaut, Johannes Twiefel

Recurrent Neural Network Sentence Parser for Multiple Languages with Flexible Meaning Representations for Home Scenarios

Xavier Hinaut[1] and Johannes Twiefel[2]

***Abstract*— We present a Recurrent Neural Network (RNN), namely an Echo State Network (ESN), that performs sentence comprehension and can be used for Human-Robot Interaction (HRI). The RNN is trained to map sentence structures to meanings (i.e. predicates). We have previously shown that this ESN is able to generalize to unknown sentence structures in English and French. The flexibility of the predicates it can learn to produce enables one to use the model to explore language acquisition in a developmental approach. This RNN has been encapsulated in a ROS module which enables one to use it in a cognitive robotic architecture. Here, for the first time, we show that it can be trained to learn to parse sentences related to home scenarios with higly flexible predicate representations and variable sentence structures. Moreover we apply it to various languages, including some languages that were never tried with the architecture before, namely German and Spanish. We conclude that the representations are not limited to predicates, other type of representations can be used.**

I. INTRODUCTION

For humans, communicating with current robots is challenging. It involves either learning a programming language or using complex interfaces. The most sophisticated ones, that recognise command given orally, are limited to a pre-programmed set of stereotypical sentences such as "Give me cup". In this paper, we propose an approach that allows one to use natural language when interacting with robots. Our architecture enables one to train a sentence parser easily on potentially many different contexts, including home scenarios (e.g. grasping remote objects or cleaning furnitures or rooms) [1]. It can also enable the users to directly teach new sentences to the robot [2].

Current approaches typically involve developing methods specifically aimed at — or only tested on — the English language. Specific parsers then need to be designed for other languages, often adapting the algorithms and methods to their specificities. Here, we propose to overcome these limitations by providing a simple and flexible way of training a sentence parser for any arbitrary language, without requiring any new programming effort. This enables to easily create parsers even for languages that are spoken by a small population, thus broadening the potential application of robotics to new communities.

*This work was partly supported by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme: EchoRob project (PIEF-GA-2013-627156).

[1]X. Hinaut is with Inria Bordeaux Sud-Ouest, Talence, France; LaBRI, UMR 5800, CNRS, Bordeaux INP, Université de Bordeaux, Talence, France; and Institut des Maladies Neurodégénératives, UMR 5293, CNRS, Université de Bordeaux, Bordeaux, France xavier.hinaut@inria.fr

[2]J. Twiefel is with Knowledge Technology Group, Department Informatik, Universität Hamburg, 22527 Hamburg, Germany twiefel@informatik.uni-hamburg.de

Here, we propose to overcome these limitations by providing a simple and flexible way of training a sentence parser for potentially any language. We demonstrate the potential of our model in home scenarios for four common European languages: English, German, Spanish and French. To make home robotics possible and useful, we need robots that are able to provide social adaptation to the user, and able to adapt continuously over their whole lifetime. Our architecture is inspired from neurosciences [3] and language acquisition theories [4][5][6] with the goal of modelling human language use: thus it aims to have these social and life-long adaptabilities.

First, we present the general ESN architecture. Then, we detail the reservoir sentence processing model. Afterwards, we show how it can be used within a ROS module, and give some perspectives. Finally, we provide corpora examples for a home scenario in four languages, and show how flexible the meaning representations can be.

II. METHODS & RESULTS

A. Echo State Networks

The language module is based on an ESN [7] with leaky neurons: inputs are projected to a random recurrent layer and a linear output layer (called "read-out") is modified by learning (which can also be done in an online fashion). The units of the recurrent neural network have a *leak rate* (α) hyper-parameter which corresponds to the inverse of a time constant. These equations define the update of the ESN:

$$\mathbf{x}(t+1) = (1-\alpha)\mathbf{x}(t) + \alpha f(W^{in}\mathbf{u}(t+1) + W\mathbf{x}(t)) \quad (1)$$

$$\mathbf{y}(t) = W^{out}\mathbf{x}(t) \quad (2)$$

with $\mathbf{x}(t)$, $\mathbf{u}(t)$ and $\mathbf{y}(t)$ the reservoir (i.e. hidden) state, the input and the read-out (i.e. output) states respectively at time t, α the *leak rate*, W, W^{in} and W^{out} the reservoir, the input and the output matrices respectively and f the *tanh* activation function. After the collection of all reservoir states the following equation defines how the read-out (i.e. output) weights are trained:

$$W^{out} = Y^d[1; X]^+ \quad (3)$$

with Y^d the concatenation of the desired outputs, X the concatenation of the reservoir states (over all time steps for all train sentences) and M^+ the Moore-Penrose pseudo-inverse of matrix M. Hyper-parameters that can be used for this task are the following: *spectral radius*: 1, *input scaling*: 0.75, *leak rate*: 0.17, *number of reservoir units*: 100.

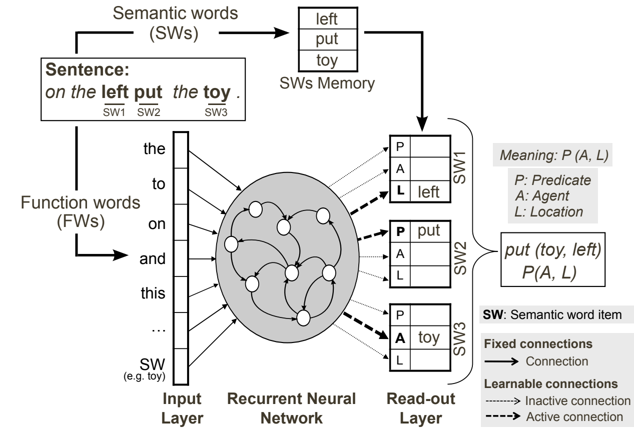

Fig. 1. Sentences are converted to a sentence structure by replacing semantic words by a SW marker. The ESN is given the sentence structure word by word. Each word activates a different input unit. During training, the connections to the readout layer are modified to learn the mapping between the sentence structure and the arguments of the predicates. When a sentence is tested, the most active units are bound with the SW kept in the SWs memory to form the resulting predicate. (Adapt. from [2].)

B. *Reservoir Sentence Processing Model*

How do children learn language? In particular, how do they associate the structure of a sentence to its meaning? This question is linked to the more general issue: how does the brain associate sequences of symbols to internal symbolic or sub-symbolic representations? We propose a framework to understand how language is acquired based on a simple and generic neural architecture (Echo State Networks) [7] which is not hand-crafted for a particular task.

The reservoir sentence processing model has been adapted from previous experiments on a neuro-inspired model for sentence comprehension using ESN [8] and its application to HRI [2]. The model learns the mapping of the semantic words (SW; e.g. nouns, verbs) of a sentence onto the different slots (the thematic roles: e.g. action, location) of a basic event structure (e.g. *action(object, location)*). As depicted in Fig. 1, the system processes a sentence as input and generates corresponding predicates. Before being fed to the ESN, sentences are transformed into a sentence structure (or *grammatical construction*) semantic words (SW), i.e. nouns, verbs and adjectives that have to be assigned a thematic role, are replaced by the *SW* item. The processing of the grammatical construction is sequential (one word at a time) and the final estimation of the thematic roles for each SW is read-out at the end of the sentence.

By processing *constructions* [5] and not sentences *per se*, the model is able to bind a virtually unlimited number of sentences to these sentence structures. Based only on a small training corpus (a few tens of sentences), this enables the model to process future sentences with currently unknown semantic words if the sentence structures are similar. One major advantage of this neural network language module is that no parsing grammar has to be defined a priori: the system learns only from the examples given in the training corpus. Here are some input/output transformations that the language model performs:

- "Please give me the mug" → *give(mug, me)*
- "Could you clean the table with the sponge?" → *clean(table, sponge)*
- "Find the chocolate and bring it to me" → *find(chocolate), bring(chocolate, me)*

As shown, the system can robustly transform different types of sentences. Recently, we have explored how flexible our system is: we found that it can handle various kinds of representations at the same time. For instance, it allows to use nouns as main elements of a predicate, and use its arguments to fill in adjectives:

- "Search for the small pink object" → *search(object), object(small, pink)*

It also allows to process various kinds of complex sentences:

- "Bring me the newspaper which is on the table in the kitchen" → *bring(newspaper, me), newspaper(on, table), table(in, kitchen)*

C. *Preprocessing*

Preprocessing to identify semantic words can be done in two ways, one may be more convenient than the other depending on the application:

- define a list of function words (i.e. all the words that are not in the meaning representation part of the construction) and consider any other word as semantic word;
- define a list of semantic words (i.e. words that are in meaning representations).

For some language like Spanish, one may need to extract some word suffixes (or prefixes) from semantic words. For instance, "sirve -lo" is usually written as one word ("sirvelo") even if its translation to English is composed of two words ("serve it").

Currently, there was no need to use part-of-speech (POS) tagging, but it may help to add this complementary information in the input representation of words.

D. *Usage of the ROS Module*

The proposed model was encapsulated in a ROS module [9]. It is coded in the Python language and uses the *rospy* library. When running the program, the reservoir model is trained using a text file and the ROS service is initialized. The training text file contains sentences and corresponding predicates and can be edited easily. If predicates represent multiple actions that have to be performed in a particular order, the predicates have to be specified in chronological order:

- `bring coffee me, clean table; bring me the coffee before you clean the table`
- `bring coffee me, cleaning table; before cleaning the table bring me the coffee`

This predicate representation enables one to easily integrate this model into a robotic architecture. It enables the users to define which kind of predicate representation they want to use. Moreover, parentheses were removed in order to make the typing and editing quicker: by convention the first word is the predicate, and the following words are the arguments.

Once initialized, a request could be sent to the ROS service: it accepts a sentence (text string) as input and returns an array of predicates in real time. With an ESN of 100 units, the training of 200 sentences takes about one second on a laptop computer. Testing a sentence is of the order of 10 ms.

E. Perspectives

In Hinaut et al. [10], it has been shown that the model can learn to process sentences with out-of-vocabulary words. Moreover, we demonstrated that it can generalize to unknown constructions in both French and English at the same time. To illustrate how the robot interaction works, a video can be seen at `youtu.be/FpYDco3ZgkU` [11][12]. The source code, implemented as a ROS module, is available at `github.com/neuronalX/EchoRob`[1].

This ROS module could be employed to process various hypotheses generated by a speech recognition system (like in [12]), then returning the retrieved predicates for each hypothesis, thus, enabling a semantic analyser or world simulator to choose the predicates with the highest likelihood. Preliminary work has shown that the model could be trained fully incrementally [13]: we plan to add this feature to the ROS module in the future.

In a nutshell, the objectives of this model are to improve HRI and provide models of language acquisition. From the HRI point of view, the aim of using this neural network-based model is (1) to gain adaptability because the system is trained on corpus examples (no need to predefine a parser for each language), (2) to be able to process natural language sentences instead of stereotypical sentences (i.e. “put cup left”), and (3) to be able to generalize to unknown sentence structures (not in the training data set). Moreover, this model is quite flexible when changing the output predicate representations, as we have shown here. From the computational neuroscience and developmental robotics point of view, the aim of this architecture is to model and test hypotheses about child learning processes of language acquisition [4].

III. Home Corpora in Multiple Languages

A. Introduction

In this section, we show that the network is able to learn to parse sentences related to home scenarios. In particular, similar networks (with the same hyper-parameters) can learn to map sentences to predicates in English, German, Spanish and French.

Each line contains a sentence together with its corresponding meaning representation: left part of the semi-column is the meaning, right part is the sentence. One can see that the way to write predicates is fairly intuitive and can be done without prior knowledge of a predefined structure. Hereafter, we show the four home corpora.

[1]We would be very grateful if you could share you corpus by uploading it to the repository. Thus we could enhance the model based on various corpora.

B. English

1) open door ; open the door
2) answer phone ; answer the phone
3) water plants ; water the plants
4) clear table ; clear the table
5) take coffee, pour coffee into mug ; take the coffee and pour it into the mug
6) clean table, put mug on table ; clean the table and put the mug on it
7) put mug on left ; put the mug on the left
8) get phone, bring phone me ; get the phone and bring it to me
9) go to bathroom ; go to the bathroom
10) make tea me; make me some tea
11) tell joke me ; tell me a joke
12) make sandwhich me, sandwhich tomatoes ; make me a sandwhich with tomatoes
13) bring newspaper me, newspaper on table, table in kitchen ; bring me the newspaper which is on the table in the kitchen
14) bring dress me, dress blue, dress in closet ; bring me the blue dress that is in the closet
15) bring pen me, pen blue, pen beside cup, cup red ; bring me the blue pen beside the red cup
16) dance for me, clean floor, clean same time ; dance for me and clean the floor at the same time
17) switch off light ; switch off the light
18) find bottle, bottle water ; find the bottle of water
19) search object, object small pink ; search for the small pink object
20) search recipe internet, recipe tiramisu ; search the recipe of tiramisu on internet
21) check if, if husband home, husband my, if five ; check if my husband is at home at five

C. German

1) öffne Tür ; öffne die Tür
2) geh an Telefon ; geh an das Telefon
3) gieße Pflanzen ; gieße die Pflanzen
4) leere Tisch ; leere den Tisch
5) nimm Kaffee, gieß Kaffee in Tasse ; nimm den Kaffee and gieß ihn in die Tasse
6) reinige Tisch, stell Tasse auf Tisch ; reinige den Tisch und stell die Tasse auf ihn
7) stell Tasse nach links ; stell die Tasse nach links
8) nimm Telefon, bring Telefon mir ; nimm das Telefon und bring es zu mir
9) geh in Badezimmer ; geh in das Badezimmer
10) mach Tee mir ; mach mir etwas Tee
11) erzähl Witz mir ; erzähl mir einen Witz
12) mach Sandwich mir, Sandwich Tomaten ; mach mir ein Sandwich mit Tomaten
13) bring Zeitung mir, Zeitung auf Tisch ; bring mir die Zeitung, die auf dem Tisch ist

14) bring Kleid, Kleid blaue, Kleid im Schrank ; bring mir das blaue Kleid , das im Schrank ist
15) bring Stift mir, Stift blauen, Stift neben Becher, Becher roten ; bring mir den blauen Stift neben dem roten Becher
16) tanz für mich, reinige Boden, reinige gleichen Zeit ; tanz für mich und reinige den Boden zur gleichen Zeit
17) schalte aus Licht ; schalte das Licht aus
18) finde Flasche, Flasche Wasser ; finde die Flasche mit Wasser
19) such Objekt, Objekt kleine pinke ; such das kleine pinke Objekt
20) such Rezept Internet, Rezept Tiramisu ; such das Rezept für Tiramisu im Internet
21) überprüfe ob, ob Mann Hause, Mann mein, ob fünf ; überprüfe , ob mein Mann um fünf zu Hause ist

D. Spanish

1) abre puerta ; abre la puerta
2) contesta teléfono ; contesta el teléfono
3) riega plantas ; riega las plantas
4) liempia mesa ; limpia la mesa
5) toma café, sirve café en taza ; toma el café y sirve -lo en la taza
6) limpia mesa, pon taza sobre mesa ; limpia la mesa y pon la taza sobre ella
7) pon taza en izquierda ; pon la taza en la izquierda
8) toma teléfono, trae teléfono -me ; toma el teléfono y trae -me -lo
9) ve baño ; ve al baño
10) prepara té -me; prepara -me té
11) di broma -me ; di -me una broma
12) prepara sandwich -me, sandwich tomates ; prepara -me un sandwich con tomates
13) trae periodico -me, periodico sobre mesa, mesa en cocina ; trae -me el periódico que está sobre la mesa en la cocina
14) trae vestido -me, vestido azul, vestido en armario ; trae -me el vestido azul que está en el armario
15) trae lápiz -me, lápiz azul, lápiz junto taza, taza roja ; trae -me el lápiz azul junto a la taza roja
16) baila para mi, limpia piso, limpia mismo tiempo ; baila para mi y limpia el piso al mismo tiempo
17) apaga luz ; apaga la luz
18) encuentra botella, botella agua ; encuentra la botella de agua
19) busca objeto, objeto rosado pequeño ; busca el objeto rosado pequeño
20) busca receta internet, receta tiramisú ; busca la receta de tiramisú en internet
21) chequea si, si marido casa, marido mi, si cinco ; chequea si mi marido está en casa a las cinco

E. French

1) ouvre porte ; ouvre la porte
2) réponds téléphone ; réponds au téléphone
3) arrose plantes ; arrose les plantes
4) débarrasse table ; débarrasse la table
5) prends café, verse café tasse ; prends le café et verse le dans la tasse
6) nettoie table, met tasse dessus table ; nettoie la table et met la tasse dessus
7) met tasse gauche ; met la tasse à gauche
8) prends téléphone, amène téléphone moi ; prends le téléphone et amène le moi
9) va dans salle_de_bain ; va dans la salle_de_bain
10) fais thé moi ; fais moi du thé
11) raconte blague moi ; raconte moi une blague
12) fais sandwich moi, sandwhich tomates ; fais moi un sandwhich avec des tomates
13) apporte journal moi, journal sur table, table dans cuisine ; apporte moi le journal qui est sur la table dans la cuisine
14) apporte robe moi, robe bleu, robe dans placard ; apporte moi la robe bleue qui est dans le placard
15) apporte stylo moi, stylo bleu, stylo côté tasse, tasse rouge ; apporte moi le stylo bleu à côté de la tasse rouge
16) danse pour moi, nettoie sol, nettoie même temps; danse pour moi et nettoie le sol en même temps
17) éteins lumière ; éteins la lumière
18) trouve bouteille, bouteille eau ; trouve la bouteille d' eau
19) cherche objet, objet petit rose ; cherche le petit objet rose
20) cherche recette internet, recette tiramisu ; cherche la recette du tiramisu sur internet
21) vérifie si, si mari maison, mari mon, si heures cinq ; vérifie si mon mari est à la maison à cinq heures

F. Remarks

All sentences in German, Spanish and French correspond to the translation from English sentences, but there are some specificities in each language which make the predicates not a direct translation word by word. For instance, in the German sentence "geh an Telefon ; geh an das Telefon" ("answer phone ; answer the phone"), the "an" is not present in the English predicate because it is a "verb particle" (a German specificity). In a similar way, for the Spanish sentence "toma el telefono y trae -me -lo" ("get the phone and bring it to me"): the direct translation of "to me" would be "a mi", but it is more natural in Spanish to attach "-me" to the verb. When *grammatical markers* such as "-me" are used in the meaning representation, we use "-" to indicate it is not a word by itself, and the direction indicates to which word this marker should be attached.

G. Results

A network of 100 neurons is able to learn a set of 21 sentence-predicate pairs in any of these languages (i.e. reproduce exactly the same predicates for each sentence). In particular, a network of 75 units is able to learn separately any of English, French or Spanish corpora. Only the German corpus needs a 100 unit reservoir to be learned perfectly[2]. The difference of units needed is not surprising since some language specificities (for this tiny highly variable data set) may be more difficult than others to learn[3]. Moreover, we found for instance that a network with just 150 neurons was able to learn both English and German corpora at the same time. A network of size 325 was needed to learn the four corpora at the same time. Thus, the learning capabilities of the network seems to increase linearly when adding new languages[4]. This needs to be confirmed with further experiments.

We did not investigate generalization concerning these corpora, even if we already demonstrated that the same network could generalize on both English and French at the same time [10]. It was out of scope here; we focus on the flexibility of the meaning representations, and in particular in home scenarios. Furthermore, the corpora are of tiny size (only 21 sentences for each language) with a high variability in their structure. Consequently, each sentence is nearly unique in its structure (and its use of function words), which does not enable generalization as shown in previous studies.

IV. Discussion

We proposed an approach that allows people to use natural language when interacting with robots. Our architecture enables to train a sentence parser easily on home scenarios: no parsing grammar has to be defined a priori, it only relies on the particular corpus used for training. Moreover, we showed that this neural parser is not limited to one language, but was able to learn corpora in four different languages at the same time: English, German, Spanish and French. In [3] an anterior version of the model was able to learn a Japanese copora, thus this neural model is applicable to many languages.

We used an *Echo State Network* (ESN) for the core part of the module for two main reasons. Firstly, it aims at modelling brain processes of language comprehension and language acquisition [8]. Reservoir Computing paradigm (such as ESNs) is considered to be a biologically plausible model of "canonical circuits" (i.e. generic pieces of cortex) [14][15] and is also used to decode neural activity of primate prefrontal cortex [16]. Secondly, we believe that HRI applications need modules that can be trained quickly (e.g. via one-shot offline learning), executed in tens of milliseconds, and can process (and provide outputs) in an incremental fashion (word by word).

The main drawback of the architecture is the fact that it only relies on function words to interpret a sentence: this may cause some ambiguities when an identical sentence structure corresponds to different meanings (see [8] for more details). However, reliance on function words is also its strength since it enables to generalize to new inputs with only tens of sentence structures in the training corpus. We believe this is not a limitation of the architecture itself since some current unpublished experiments show that real semantic words could be used instead of SW markers (i.e. fillers).

The ability of the model to scale to larger copora depends on the similarity of the sentence structures in the corpus. If sentence structures are different variants of similar expressions (usually in a same home context), then there is no need to use an high number of units in the reservoir. For instance in [8] we have shown that a 5.000 units reservoir could generalize up to 85% on unkown sentence structures on a corpus of size 90.000. Conversely, if sentence structures are highly different, the number of units should be increased (apparently linearly) with the number of sentences. Moreover, if sentence structures are similar or applied to the same home context (e.g. kitchen, living room, ...), the generalization to new sentence structures will come for free (with a quick tuning of main parameters: spectral radius, leak rate and input scaling). In future work, we could limit the number of reservoir unit needed if some contextual inputs are provided to the network: for instance, kitchen or living room environment will not provide the same context input, thus giving the network more information to discriminate/parse input sentences.

The way to write predicates (left hand of each line) is fairly intuitive and can be done without prior knowledge of a predefined structure like Robot Control Language (RCL) commands [17]. In other words, we propose to let the users define which meaning/predicate representation they want to use. For instance, some words like "on" can be used in the meaning representation when needed (e.g. "clean table, put mug on table ; clean the table and put the mug on it"), and can be discarded when they are not necessary (e.g. "search recipe internet, recipe tiramisu ; search the recipe of tiramisu on internet"). Accordingly, one can adapt the meaning representation to any kind of "series of slots" as soon as there are consistent with each other: consistency should be kept both within slots and in the linear organization of the slots. In conclusion, this means that we are not limited to predicates, and our architecture could be used to learn other type of representations.

Acknowledgment

We thank Ikram Chraibi Kaadoud, Pramod Kaushik, Louis Devers, and Iris Wieser for their ideas on the home corpora, Francisco Cruz for his nice help on the Spanish corpus, and Fabien Benureau for his very/damn useful feedback.

References

[1] F. Cruz, J. Twiefel, S. Magg, C. Weber, and S. Wermter. Interactive reinforcement learning through speech guidance in a domestic scenario. In *Proc. of IJCNN*, pages 1–8. IEEE, 2015.

[2] X. Hinaut, M. Petit, G. Pointeau, and P. F. Dominey. Exploring the acquisition and production of grammatical constructions through human-robot interaction with echo state networks. *Frontiers in Neurorobotics*, 8, 2014.

[3] P.F. Dominey, M. Hoen, and T. Inui. A neurolinguistic model of grammatical construction processing. *Journal of Cognitive Neuroscience*, 18(12):2088–2107, 2006.

[2] We only changed the network size by 25 unit steps.

[3] In our experiments, the German sentence-predicate pair "switch off light; switch off the light" was the only one not learned by a 75 units network.

[4] $75(English)+75(Spanish)+75(French)+100(German) = 325$

[4] M. Tomasello. *Constructing a language: A usage based approach to language acquisition*. Cambridge, MA: Harvard University Press, 2003.

[5] A.E. Goldberg. *Constructions: A construction grammar approach to argument structure*. University of Chicago Press, 1995.

[6] E. Bates, B. Wulfeck, and B. MacWhinney. Cross-linguistic research in aphasia: An overview. *Brain and language*, 41(2):123–148, 1991.

[7] H. Jaeger. The “echo state” approach to analysing and training recurrent neural networks. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148:34, 2001.

[8] X. Hinaut and P.F. Dominey. Real-time parallel processing of grammatical structure in the fronto-striatal system: a recurrent network simulation study using reservoir computing. *PloS one*, 8(2):e52946, 2013.

[9] X. Hinaut, J. Twiefel, and S. Wermter. Recurrent neural network for syntax learning with flexible predicates for robotic architectures. In *Proc. of the IEEE Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. IEEE, 2016.

[10] X. Hinaut, J. Twiefel, M. Petit, P. Dominey, and S. Wermter. A recurrent neural network for multiple language acquisition: Starting with english and french. In *NIPS 2015 Workshop on Cognitive Computation: Integrating Neural and Symbolic Approaches*, 2015.

[11] X. Hinaut, J. Twiefel, M. Borghetti Soares, P. Barros, L. Mici, and S. Wermter. Humanoidly speaking – learning about the world and language with a humanoid friendly robot. In *IJCAI Video competition, Buenos Aires, Argentina*. https://youtu.be/FpYDco3ZgkU, 2015.

[12] J. Twiefel, X. Hinaut, M. Borghetti, E. Strahl, and S. Wermter. Using Natural Language Feedback in a Neuro-inspired Integrated Multimodal Robotic Architecture. In *Proc. of RO-MAN*, New York City, USA, 2016.

[13] X. Hinaut and S. Wermter. An incremental approach to language acquisition: Thematic role assignment with echo state networks. In *Proc. of ICANN 2014*, pages 33–40, 2014.

[14] P.F. Dominey. Complex sensory-motor sequence learning based on recurrent state representation and reinforcement learning. *Biological cybernetics*, 73(3):265–274, 1995.

[15] W. Maass, T. Natschläger, and H. Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural computation*, 14(11):2531–2560, 2002.

[16] M. Rigotti, O. Barak, M.R. Warden, X.J. Wang, N.D. Daw, E.K. Miller, and S. Fusi. The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451):585–590, 2013.

[17] K. Dukes. Semeval-2014 task 6: Supervised semantic parsing of robotic spatial commands. *SemEval 2014*, page 45, 2014.