Natural Noun Phrase Variation for Interactive Characters

Donna K. Byron, Aakash Dalwani, Ryan Gerritsen, Mark Keck
Thomas Mampilly, Vinay Sharma, Laura Stoia, Timothy Weale and **Tianfang Xu**
The Ohio State University Computer Science Department
2015 Neil Avenue
Columbus, Ohio USA 43221

Abstract

Interactive characters that cohabit a shared space with human partners need to generate and interpret references to elements of the virtual world. Natural language allows for a wide range of phrasings for referring to any particular object, and this variation is thought to reflect not only spatial but also cognitive and linguistic factors. Our study attempts to account for the variability in referring forms found in a set of dialogs of two human partners performing a treasure-hunt task in a virtual world. A decision tree classifier was built that predicts the form of 51% of the referring expressions, compared to a baseline of 39% achieved by a heuristic classifier. The classification algorithm can be used by conversational characters to generate referring expressions of the appropriate form.

Introduction

Natural language interaction with simulated characters has been a long standing goal of both game developers and artificial intelligence researchers. Rudimentary language processing is now included in commercially available computer games, such as Lifeline from Konami[1], which allows the user to speak short commands to the game character such as "Load" and "Shoot". Natural language capabilities have been added into research prototype systems such as the Nice fairy tale system (Gustafson *et al.* 2004) and text-based adventure games (Gabsdil, Koller, & Striegnitz 2002). Full natural interaction, allowing the user to express himself in the richly varied language of everyday conversation, will enable the user to employ a rich set of linguistic inputs, but this increases the likelihood that his input will be ambiguous. One of the most troublesome points of ambiguity in natural language comes from definite noun phrases, which provide the arguments to commands. For example, "Let's see what's in that room" might be spoken as a command to a game character to explore a particular room, but the character must determine which room it has just been instructed to go into. Richer linguistic interaction of this type is needed for natural spoken interaction with animated agents in many domains, not only games, but also interactive social agents and storytelling partners, agents in training simulations, and robotic agents performing a task with human partners.

[1] http://www.konami.com/gs/lifeline/

Due to their complexity, nominal referring expressions (such as "that room" in the example above) are typically interpreted in language processing software by a dedicated module. For example, in the TRIPS spoken dialog system (Ferguson & Allen 1998), the *Reference Manager* determines the meaning of referring expressions spoken by the user and helps the system choose an appropriate phrasing when it speaks back to the user (Allen *et al.* 2000). In the Galaxy architecture developed at MIT (Seneff *et al.* 1998), a separate *Context Manager* module performs this task (Filisko 2002). An augmented reality system developed at OGI allows the user, positioned in real space, to point to virtual objects registered into his field of view through a graphical overlay. This system once again makes use of a dedicated module to understand the user's references (Kaiser *et al.* November 2003). Reference processing is a challenging task because references are formulated in context, and the exact factors that modulate the context's effect on language are not yet well understood.

This paper concentrates on the particular difficulties of building reference processing for virtual partners (bots) that are visualized as part of a 3D graphical world and can discuss the world they are in and the tasks to be performed there with a human partner. An important property of such situations is that both the bot and the human partner are represented by avatars which are localized within the task world and can move about independently within the world. The contextual effects on reference processing for such situations is different enough from traditional computational linguistics paradigms to present difficult challenges for algorithm building. This difficulty arises from the fact that when a human and a virtual partner carry on a dialog while they are situated within a 3D space, the form and focus of their conversation at each time step can be influenced by three factors:

1. The task they are performing in the virtual world
2. The elements of the virtual world they can see (or have seen), and their spatial configuration
3. The form and content of the conversation-so-far

Because noun phrases (NPs) are highly context sensitive, the interaction of these three factors must be determined before naturalistic NP processing can be added to conversational agents in virtual worlds. As a preliminary to building reference processing, this paper examines how speakers in 3D dialog utilize different forms of referring expressions to describe objects in their world. The goal is to discover regularities in the properties of the objects described and the linguistic variation employed by human speakers. Any such regularities can be utilized in bots for both reference understanding software and also to produce appropriate, natural sounding NPs by virtual characters with natural language generation capabilities.

Background

When a speaker refers to a particular entity in the world, a variety of different linguistic forms might be used for the referring expression. For example, a particular building can be called “my house” (a description), “42 Main Street” (a name), “that eyesore” (a demonstrative description) or “it” (a pronoun), depending on the situation and the person being addressed. Although these different forms can all refer to the same referent, at a specific point in a discourse, the speaker is not completely at liberty to employ any alternative at all. Providing too much or too little information is perceived as uncooperative (Grice 1975). Psycholinguistic studies indicate that using the wrong form at the wrong time can make a sentence confusing or even incomprehensible (Garnham 2001).

As an engineering decision, a conversational bot might be designed to only understand and produce particular forms, for example, a common strategy is to assign every object in the world a unique designator that must be used for that item (as in example (1)). This strategy has multiple disadvantages. Requiring the human user to speak in an unnatural fashion will force him to concentrate on formulating his sentences properly and distract him from attending to the task at hand. When the same item is repeated in a discourse, it is more natural to use a reduced form, as the second line of example (2) demonstrates. If the bot produces an inappropriate form, such as in the fourth sentence of example (1) when the bot repeats “the control room” rather than a reduced form, psycholinguistic studies indicate that this is likely to confuse the human partner (Chambers *et al.* 2002). Also, it is awkward to teach the human player the correct names for objects in the world, especially for objects that do not naturally take on names, such as staircases, hairbrushes, or corners of rooms. For such objects, the most desirable strategy is to let the user refer to them with a natural phrasing of his own choosing, and for the bot to work out what those words mean. Of course, 50 years of AI research has demonstrated that this is not a simple task.

(1)

Conversation using only rigid designators	
Human:	We need to go back to **the control room** and find **the tactical officer**
Bot:	Where was **the control room**?
Human:	**The control room** is on **the second level** near **the stairs**
Bot:	Right, let’s go to **the control room**

(2)

Conversation using NP variation	
Human:	We need to go back to **the control room** and see if **the tactical officer** is **there** now
Bot:	Where was **that**?
Human:	**It** is **upstairs** just by **the staircase**
Bot:	Right, let’s go

Choosing the correct forms seems to involve a complex interdependency between properties of the speaker and hearer and properties of the referent. This background section reviews prior work that clarifies which properties should be taken into account. In the experiments described below, we will put these claims to the test on situated dialog data.

Table 1: The Givenness Hierarchy (for Definite NPs)

Status	in focus >	activated >	familiar >	uniquely identifiable
NP Form	*it/them they*	*this/that this N*	*that N*	*the N*
	(PRP)	(CDP/DP/ CDEM)	DEM	DEF

NPs can be phrased using either definite (e.g. *the house*, *my house*) or indefinite forms (e.g. *a house*). Using a definite expression signals that the speaker assumes that the object is already known to the hearer, or easily inferable from what he knows (Roberts 2003; Prince 1981). In the remainder of this paper, we will restrict our focus to the variety of forms of definite referring expressions. There are many forms of definite expression, so the high-level split of definite/indefinite is not sufficient to choose a phrasing for a referring expression. The following terminlogy will be employed for definite expression types throughout this paper:

- Descriptive Definite **(DEF)**: A common noun head, e.g. *house*, plus a definite or possessive determiner, e.g. *the/my*, or optional quantifier, e.g. *the two buttons*, as well as modifiers.
- Demonstrative Description: The determiner *this, that, these, or those* plus a descriptive head, e.g. *house* to produce e.g. *those houses*. We will distinguish Close Demonstrative Descriptions **(CDEM)** which use the determiners *this* and *these* from unmarked Demonstrative Descriptions **(DEM)** with the determiners *that* and *those* (Lyons 1977).
- Demonstrative Pronouns: The pronoun *this, that, these, those* on their own with no common noun head. Again, we distinguish between the close-marked ones **(CDP)** and unmarked ones **(DP)**.
- Personal Pronoun **(PRP)**: Reduced forms with no common noun head such as *he/she/them/it*.

This leaves us with a set of 6 labels (PRP, CDP, DP, CDEM, DEM, DEF) to distinguish definite NP forms. What controls a speaker’s choice among these forms at a particular moment in time? Theoretical work such as the Givenness

Hierarchy, shown in Table 1 (Gundel, Hedberg, & Zacharski 1993) (henceforth GH) claims that each different NP form signals the speaker's assumption about the cognitive status of a referent with the type(s) of referring expression that conventionally signal that status. The GH states that DEF signals *uniquely identifiable*, which means that the addressee can attach a unique entity to the referring expression after hearing the full sentence, whether or not he had previous knowledge of that entity. This category is distinguished from *familiar*, in which the addressee is expected to have prior knowledge of the item being mentioned. *Activated* status, used for DEM, CDEM, and CDP requires that the referent be already in the addressee's short term memory, having been evoked by either the discourse or the setting.

What evidence can be used to decide if one's conversational partner has prior knowledge of an item? (Clark & Marshall 1981) call this *mutual knowledge*, and define the following types of mutual knowledge:

Physical copresence:

- Current: An item that the partners are simultaneously attending to
- Prior: An item seen before, that is memorable
- Potential: An item not seen yet, that is locatable

Linguistic copresence: An item that was discussed before and that is memorable.

Finally, referents of PRPs are claimed to have *in Focus* status. The GH model defines focus as the: "partially-ordered subset of activated entities that are likely to be continued as topics of subsequent utterances"[pg. 279]. Computational models often use surface linguistic clues to estimate which item will be the topic of a subsequent utterance. The grammatical role of each NP in a sentence is a commonly used surface clue. In English, there is a high probability that the top-level arguments of the verb will appear in focus in the next sentence. In the sentence "Bruce put those things next to the table", *Bruce* is in subject position, and *those things* is the direct object, and the relative focus ordering on these items would be $Bruce > ThoseThings > Table$. In addition to surface order properties, other parameters such as how many times a particular item has been mentioned in the discourse can be used to estimate topicality, and these features have been found to be very useful for matching PRPs to their referents (see (Mitkov 2002) for a recent survey). Focus is acknowledged in the literature to be controlled by a variety of factors beyond linguistic clues, but these are most commonly used since they can be reliably computed.

The task that the partners are performing also places the focus on items important to the task (Grosz 1977; Kasper, Davis, & Roberts 1999). In our treasure-hunt domain, once the partners decide that they will perform a particular objective next, such as finding a helmet, that item remains focused until the task objective shifts. In addition, the items in the scene, especially items that the speaker is currently looking at, are also likely to be continued as topic.

Taken together, these claims about the factors that modulate a speaker's choice of NP form seem to converge around three significant dimensions:

Figure 1: One player's view when he says: "the things the table we need to move those things around"

Spatial The arrangement in 3D space of the speaker, hearer, and object. The important distinction in this dimension can be called COPRESENCE: the item is now perceptually accessible to both the speaker and hearer at the point when the NP is spoken, or it was COPRESENT earlier.

Cognitive Whether the addressee already knows about the object, or if it is brand new to him. If it is already known, to what extent is his attention focused on the object. The relevant categories here are: FOCUS, ACTIVATED, FAMILIAR.

LINGUISTIC Has the object been mentioned before, how recently, and in what argument position.

Situated Interaction in Quake

The data studied here was collected as part of a pilot project underway in the SLATE lab at OSU[2]. In this study, two human partners were established as 'players' in a first-person graphical world, rendered by the QuakeII game engine[3]. The partners are in the virtual world, where they collaborate on a treasure hunt task working from printed instructions that show pictures of several items they must either manipulate or find and move to a new location. Both players have the same capabilities for performing the task. The partners communicate with each other through headset-mounted microphones, and an audio recording of their dialog was saved and later transcribed. In addition, each player's movement and activity in the virtual world was recorded to video tape.

The QuakeII game engine allows the two partners to move about in the world independently and manipulate objects. As he moves about, each player sees a first-person view of the virtual world. When both players position themselves so that they have an unobstructed view of each other, they can observe the avatar representing the other player and its actions in the virtual world. Therefore, the two players can stand in the same virtual setting and discuss what they see or what

[2]http://slate.cis.ohio-state.edu

[3]available as open source from www.id.com

88-4:	A:	cause I'm gonna drop it
89-1:	B:	yeah
89-2:	B:	there it is
90-1:	A:	yeah
90-2:	A:	so that needs to go there
90-3:	A:	and then we just need a box to appear
91-1:	B:	is it one of the metal boxes
92-1:	A:	yeah
92-2:	A:	one of the metal the cubicle boxes with the
92-3:	A:	I guess they're metal
92-4:	A:	the brown ones brownish grey
93-1:	B:	is is it one of the items that we can pick up
94-1:	A:	no

Figure 2: Actual treasure hunt dialog fragment

they plan to do, and can observe their partner's movements in the world, as shown in Figure . When they are not together in the world, they tend to inform each other of their position. As a result, when the players move about the world independently, they can still mentally estimate which items in the world their partner knows about and which items he is not yet familiar with.

The human partners use a variety of NP forms to refer to objects in the virtual world. Figure 2 shows a small excerpt, in which a particular item is referred to with *it* in 88-4 and *that* in 90-2, and the metal box they are searching for goes from *a box* to *it*. It should be noted that this human-human data represents an extreme case of unconstrained language, which is likely to differ from how humans would speak to a bot for the same task (Jönsson & Dahlbäck 1988; Oviatt 1985; Doran *et al.* 2001). However, this data provides a rich testbed for understanding NP forms, and is useful especially for developing algorithms for language generation software in automated characters.

Looking at the variety of expressions found in this corpus, the question that naturally arises is:

1. How do the spatial and cognitive properties of entities in the world correlate to the different referring forms? If a bot were trying to participate in this conversation, how should it choose whether to call the box *it* or *the box*, for instance.

2. Given a referring expression to interpret, can a bot accurately infer the spatial and cognitive properties of the object being referred to, and use this information to interpret the expression?

Experiments

These experiments were undertaken to discover the extent to which variation in the NP forms found in the treasure hunt dialogs could be accounted for using the properties described above in Section 2. This section compares a heuristically constructed classifier, using hand-constructed rules to apply the claims from previous authors, to a classifier induced using machine learning. The test data used in these experiments comes from transcribing three pilot human-human task sessions involving 6 distinct speakers. The transcripts represent 70 minutes of elapsed time and contain a total of 10,547 words, with 987 definite NPs (Table 2).

NP Form	Frequency
Lexical NPs	
The N (DEF)	383
This N (CDEM)	90
That N (DEM)	68
Pronouns	
This (CDP)	59
That (DP)	130
It/them/they (PRP)	257
Total	987

Table 2: Number of test cases in the dataset

Each noun phrase NP_i listed in Table 2 was assigned one of the six labels (PRP, CDP, DP, CDEM, DEM, DEF), and was annotated with the following attributes:

TASKFOCUS The object that has been established as the currently active goal of the task. Referents outside of the task world are labeled MOOT.

LINGFOCUS The highest ranked item mentioned in the sentence before the sentence containing NP_i (using grammatical role rankings) is the linguistic focus at the start of NP_i. If NP_i mentions that same item, it is marked with LINGFOCUS=Y, otherwise LINGFOCUS=N

LINGCOPRES The item has been mentioned in the dialog at some point before NP_i is spoken. LINGFOCUS will necessarily be a subset of LINGCOPRES.

PHYSCOPRES Using Clark and Marshall's terms, each item in the world is labeled according to its perceptual accessibility. Possible values are:

MOOT = the item is un-see-able (ex. generic concepts, facts, etc.)

NO = the item has never been seen

PRIOR = the item was seen before but is not in view now

CURRENT = the players can see the object now

POTENTIAL = the item might be seen if the players look around at their current position

Baseline Classifier

Based on the background material described above in Section 2, we expect the following properties to hold for the NPs in this corpus. We constructed a hand-crafted classifier that used these rules to label the test cases:

PRP Is expected to be used for items with linguistically-established focus or task focus

CDP, CDEM Are expected to be used for items in the scene but not in focus

DP Should correlate with high-order items that cannot have physical copresence (e.g. "let's do that"), or items that the speaker wishes to draw the hearer's attention to (e.g. "what's THAT?").

DEM are not in the scene but known and discussed previously

88-4:	A:	cause I'm gonna drop **it**
89-1:	B:	yeah
89-2:	B:	there **it** is
90-1:	A:	yeah
90-2:	A:	so **IT** needs to go there
90-3:	A:	and then we just need a box to appear
91-1:	B:	is **it** one of **the metal boxes**
92-1:	A:	yeah
92-2:	A:	one of the metal **the cubicle boxes** with the
92-3:	A:	I guess **THE BOXES** are metal
92-4:	A:	**the brown ones brownish grey**
93-1:	B:	is is **it** one of **the items** that we can pick up
94-1:	A:	no

Figure 3: Dialog fragment using the Decision Tree rules to choose the form of NPs. UPPERCASE NPs in 90-2 and 92-3 differ from the original dialog.

DEF items not in the current setting, and either seen before but not discussed, or unseen

Therefore, the logic for classifying each test item is:

```
case LINGFOCUS
Y: form=PRP;
N: case TASKFOCUS
    Y: form=PRP;
    N: case PHYSCOPRES
          MOOT: form=DP;
          PRIOR: case LINGCOPRES
                    Y: form=DEM;
                    N: form=DEF;
          CURRENT: form=oneof(CDP|CDEM);
          POTENTIAL: form=DP;
          NO: form=DEF;
```

These criteria correctly accounted for 39% of the distribution of forms in this corpus. Table 3 shows the performance on each NP form for the baseline classifier using the standard metrics of Precision (the ratio of items from each category that received the right label) and Recall (the number of items assigned each label divided by the number that should have been assigned that label).

		Baseline			Decision Tree		
NP Form	#	Right	Prec.	Recall	Right	Prec.	Recall
DEF	383	57	.15	.73	299	.78	.53
CDEM/CP	150	77	.51	.40	0	0	0
DEM	68	5	.07	.05	4	.06	.25
DP	130	25	.19	.29	46	.35	.53
PRP	257	224	.87	.42	160	.62	.51
Total	987	388	.39	.39	509	.51	.51

Table 3: Comparison of Algorithm performance

Decision Tree Algorithm

Since there are multiple factors in the criteria for NP form, it seems appropriate to use machine learning to attempt to improve the hand-crafted classifier. We built a decision tree classifier using the WEKA toolkit (Witten & Frank 2000) to learn the decision process to assign NPFORM labels to this test data. The best-performing decision tree it induced is shown in Figure 4.

```
taskfocus = no
|   physcopres = no
|   |   lingfocus = no: DEF
|   |   lingfocus = yes: PRP
|   physcopres = curr: DEF
|   physcopres = potential: DEF
|   physcopres = moot
|   |   lingcopres = no: DEF
|   |   lingcopres = yes: DP
|   physcopres = prior: DEF
taskfocus = yes
|   physcopres = no: PRP
|   physcopres = curr
|   |   lingcopres = no: DEM
|   |   lingcopres = yes: PRP
|   physcopres = potential
|   |   lingfocus = no: DEF
|   |   lingfocus = yes: PRP
|   physcopres = moot
|   |   lingfocus = no: DP
|   |   lingfocus = yes: PRP
|   physcopres = prior
|   |   lingfocus = no: DEF
|   |   lingfocus = yes: PRP
taskfocus = moot: DEM
```

Figure 4: Decision Tree for labeling NPFORM

Decision tree classification accuracy on the test data are also shown in Table 3 in the columns marked ". Note that these results were obtained on the test data used to build the classifier. Performance on unseen data was only slightly lower (the decision tree correctly labeled 50% of cases in 3-fold cross-validation). No rules were generated in the decision tree for either of the close-marked forms THIS or THIS NP. This implies that the feature set developed thus far does not capture the right dimensions that are important determiners for using this form. Adding more fine-grained attentional features such as spatial proximity and gesture would probably improve this category.

These rules can be used to generate referring expressions of the appropriate form. Figure 3 shows the dialog fragment from Figure 2 as it would appear with the definite noun phrases replaced by the form determined by the decision tree. Note that, because of the overlapping usage conditions for some forms, there are cases where our algorithm does not produce the same form as the human speakers did, yet the form it does generate is still felicitous. For example, our algorithm produces *it* in utterance 90-2 rather than *that*. In future work, as more data is available, we plan to evaluate the generated output by asking human judges whether the wording produced by our algorithm is acceptable in context.

Summary

This study makes a first attempt at establishing the rules for wording NPs appropriate to the context in situated dialog. The rules can be used to either understand language produced by people, or for a bot working in a virtual world to produce the full variety of forms in the right circumstances. The current classifier predicts the correct label for only 51%

of our test cases, so obviously more work is needed to develop more sensitivity in the algorithm. However, we feel this represents an important first stride toward building detailed specifications of noun phrase processing for agents working in rich, multi-modal situations such as computer games and training simulations.

Acknowledgments

The authors would like to acknowledge our gratitude to Brad Mellen and David Menninger for their work creating the QuakeII data collection environment, and Whitney Gegg-Harrison for help transcribing the recordings.

References

Allen, J.; Byron, D.; Dzikovska, M.; Ferguson, G.; Galescu, L.; and Stent, A. 2000. An architecture for a generic dialogue shell. *Natural Language Engineering* 6(3&4):213–228.

Chambers, C.; Tanenhaus, M.; Eberhard, K.; Carlson, G.; and Filip, H. 2002. Circumscribing referential domains in real time language comprehension. *Journal of Memory and Language* 47:30–49.

Clark, H. H., and Marshall, C. R. 1981. Definite reference and mutual knowledge. In Joshi, A. K.; Webber, B. L.; and Sag, I., eds., *Elements of Discourse Understanding*. Cambridge University Press. 10–62.

Doran, C.; J.Aberdeen; Damianos, L.; and L.Hirshman. 2001. Comparing several aspects of human-human and human-computer dialogues. In *Proceedings of the 2nd Sigdial workshop on discourse and dialog*.

Ferguson, G., and Allen, J. F. 1998. TRIPS: An intelligent integrated problem-solving assistant. In *Proceedings of the National Conference on Artificial Intelligence (AAAI '98)*.

Filisko, E. A. 2002. A context resolution server for the GALAXY conversational system. Master's thesis, Massachusetts Institute of Technology Department of Electrical Engineering and Computer Science.

Gabsdil, M.; Koller, A.; and Striegnitz, K. 2002. Natural Language and Inference in a Computer Game. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING02)*.

Garnham, A., ed. 2001. *Mental models and the interpretation of anaphora*. Psychology Press.

Grice, H. P. 1975. Logic and conversation. In Cole, P., and Morgan, J., eds., *Syntax and Semantics*, 41–58. New York: Academic Press.

Grosz, B. J.; Jones, K. S.; and Webber, B. L., eds. 1986. *Readings in Natural Language Processing*. Morgan Kaufmann.

Grosz, B. J. 1977. The representation and use of focus in a system for understanding dialogs. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI '77)*, 67–76. Reprinted in (Grosz, Jones, & Webber 1986).

Gundel, J. K.; Hedberg, N.; and Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2):274–307.

Gustafson, J.; Bell, L.; Boye, J.; Lindstrom, A.; and Wire, M. 2004. The nice fairy-tale game system. In *Proceedings of the 5th Sigdial Workshop on Discourse and Dialog*.

Jönsson, A., and Dahlbäck, N. 1988. Talking to a computer is not like talking to your best friend. In *Proceedings of the first Scandinavian Conference on Artificial Intelligence*, 297–307.

Kaiser, E.; Olwal, A.; McGee, D.; Benko, H.; Corradini, A.; Li, X.; Cohen, P.; and Feiner, S. November 2003. Mutual disambiguation of 3d multimodal interaction in augmented and virtual reality. In *Proceedings of the 5th international conference on Multimodal interfaces (ICMI 2003)*.

Kasper, R. T.; Davis, P. C.; and Roberts, C. 1999. An integrated approach to reference and presupposition resolution. In *The Relation of Discourse/Dialogue Structure and Reference: Workshop at the 37th Annual Meeting of the Association for Computational Linguistics (ACL-99)*, 1–10.

Lyons, J. 1977. *Semantics*. Cambridge University Press.

Mitkov, R. 2002. *Anaphora Resolution*. Longman.

Oviatt, S. L. 1985. Predicting spoken disfluencies during human-computer interaction. *Computer Speech and Language* 9(1):19–35.

Prince, E. F. 1981. Toward a taxonomy of given-new information. In Cole., ed., *Radical Pragmatics*. Academic Press. 223–255.

Roberts, C. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26(3):287–350.

Seneff, S.; Hurley, E.; Lau, R.; Pao, C.; Schmid, P.; and Zue, V. 1998. GALAXY-II: A reference architecture for conversational system development. In *Proceedings of ICSLP-98*.

Witten, I. H., and Frank, E. 2000. *Data Mining: Practical machine learning tools with Java implementations*. San Francisco: Morgan Kaufmann.