Accentuation and Interpretation

Hans-Christian Schmitz & Bernhard Schröder

Institute for Communication Research and Phonetics University of Bonn, Germany {hcs; bsh}@ikp.uni-bonn.de

Abstract

Optimal accentuation of a sentence involves accentuating a minimal set of words which in a given context suffices for understanding the entire sentence. We propose a model of the interpretation of incomplete or not entirely recognized utterances. Using this model, we determine which constituents of an utterance have to be accentuated given a certain context.

1. Optimal Accentuation

Our basic assumption is that accentuation helps to ensure the listener's recognition of certain words. These words are sufficent for interpreting the entire utterance in a given context. The recognition of other words is not necessary for understanding. We propose a model for the interpretation of not entirely recognized utterances, like acoustically disturbed utterances, and of intrinsically incomplete utterances, like constituent-answers and telegrams. Speakers want listeners to understand their utterances. They utter their sentences in a way that leads the listeners to the interpretations. Accentuation is driven by the listeners' interpretative needs. It is one means of following Grice's maxim of manner ([2]) "Be perspicuous!".

For the correct interpretation of an utterance, the listener has to be aware of its context. Knowledge of the current issue or *question under discussion (qud)* (cf. [1]) is needed in order to derive the exhaustive meaning of an utterance. Consider example 1.1:

Example 1.1

qud 1 Which professor talked to Mary?

qud 2 Which student talked to Mary?

answer Yves talked to Mary.

As an exhaustive answer to the first *qud*, the sentence "Yves talked to Mary" is interpreted in the sense of "Yves and no other professor talked to Mary". As an exhaustive answer to the second *qud*, the sentence is interpreted in the sense of "Yves and no other student talked to Mary". Clearly, the interpretations are not equivalent.

Listeners can grasp the propositional meaning of an utterance even if an incomplete sentence was uttered or the utterance was not fully recognized. The most simple example is that of a constituent answer:

Example 1.2

qud Who talked to Mary?

answer Yves.

In example 1.2, "Yves" is not a full sentence. But regarding the *qud*, the incomplete sentence can be understood as a complete answer in the sense of "Yves talked to Mary".

A *qud* does not have to be explicitly asked, but can be implicitly given. Kuppevelt ([5]) proposes to structure any text by a series of questions. As an example, consider the telegram 1.3.

Example 1.3

telegram Plane delayed. Don't wait. Take train.

Assume that the recipient is waiting for the sender at an airport. A plane is delayed, presumably the sender's plane, because information about another plane would be of no interest. Now, if the recipient assumes that the *qud* is what the sender wants to do, then he would interpret the telegram in the sense of "The sender does not wait for the plane. He will take a train instead". If, otherwise, the recipient assumes that it is under discussion what the recipient himself should do, then he would interpret the telegram in the sender, but take a train." The recipient might assume a series of *quds* as well, e.g. "What shall I do? And what will the sender do?" Accordingly, the recipient's interpretation would be: "I shall not wait for the sender. He will take a train." Example 1.3 shows that assuming the right *quds* can be crucial for correctly understanding a text.

Acoustic communication can be disturbed: Speech signals can be masked by noise and other speech signals. Moreover, speakers do not articulate each word with the same precision. Listeners recognize a word more easily when it was precisely articulated than when it was imprecisely articulated. Even if all words of an utterance were well-articulated and acoustically not disturbed, some words might not be recognized. The listener's attention varies, distractions disturb his perception. For a listener there is a certain probability of not correctly recognizing all words of a given utterance. Nevertheless, listeners do not standardly ask the speaker to repeat. They rely on having recognized the important words which lead them to proper understanding.

An accent appears in a speech signal as an abrupt change of the pitch and a relative raising of intensity (loudness) and duration (length). Abruptly changing the pitch and raising intensity and duration are the means of raising the prominence of a word. If two subsequent words are to be accentuated, then it has to be assured that the second word is made prominent with respect to the context, i.e. the first word. Data show that this can be achieved e.g. by further changing the pitch and raising the intensity. Generally, accentuated words are more precisely articulated than non-accentuated words.

Accentuated words are more likely to be recognized than non-accentuated words. The fewer words are accentuated, the higher is the probability that the listener's attention is directed to these words and that the listener correctly recognizes all accentuated words. Therefore, optimal accentuation of an utterance involves accentuating a minimal set of words which includes those words which the listener must correctly recognize in order to understand the utterance.

The set of words of an utterance which are optimally to be accentuated in a given context is determined by a model of the interpretation of not entirely recognized utterances. Constraints of optimal accentuation are pragmatically motivated. They can interfere with syntactically motivated constraints on intonation.

2. Interpretation of not entirely recognized assertions

We limit our reflections to the interpretation of assertive utterances. Assume a competent listener of English with knowledge of the current *qud*. By applying his lexical and grammatical knowledge, he can derive semantic representations of the recognized expressions. These representations can be type-shifted and applied to the *qud*. Thereby, propositional representations of entire assertions can be derived.

Let us consider example 2.1: The words set in upper case and bold face are to be accentuated. All answers can be understood even if only the accentuated words are recognized.

Example 2.1

qud 1 Which student talked to Mary?

 $\langle \lambda x[talk(x, mary)], \{x | x \in cs \land student(x)\} \rangle$

answer 1-1 YVES talked to Mary.

answer 1-2 ALL students talked to Mary.

answer 1-3 The FRENCH student talked to Mary.

qud 2 Which Frenchman talked to Mary? $\langle \lambda x [talk(x, mary)], \{x | x \in cs \land french(x)\} \rangle$

answer 2-1 The french STUDENT talked to Mary.

answer 2-2 ALL french STUDENTS talked to Mary.

qud 3 Who talked to Mary? $\langle \lambda x[talk(x, mary)], cs \rangle$

answer 3-1 The FRENCH STUDENT talked to Mary.

answer 3-2 ALL FRENCH STUDENTS talked to Mary.

We represent wh-questions as pairs of a λ -abstract and a question-domain. E.g., "What kind of student talked to Mary" is represented as $\langle \lambda P[\exists x[P(x) \land student(x) \land talk(x, mary)]], cs \rangle$ where cs is some contextually given set of properties. "Who talked to Mary?" is represented as shown in example 2.1 (qud 3), where cs is some contextually given set of persons. Which-questions are interpreted like who- or whatquestions with a further restriction of the question-domain (cf. example 2.1, qud 1,2).

To allow for contextual domain restrictions, we introduce domain variables. Therefore, we represent the meaning of an NP as a generalized quantifier, where the restrictor of the quantifier is contextually bound. E.g., we interpret "all students" as the set of properties that are shared by all students of some contextually given domain $D: \lambda P[\forall x[x \in D \land student(x) \rightarrow P(x)]].^1$

Interpretations of the answers 1-1, 2-2 and 3-2 can be derived from interpretations of the accentuated words by the application of these interpretations to the *qud* and the instantiation of the domain-variable D by the *qud*-domain. Let it be given that the listener of answer 2-2 only recognizes "all" and "students". From his lexical and grammatical knowledge, he derives the quantifier $\lambda P[\forall x[x \in D \land student(x) \rightarrow P(x)]]$ as a representation of "all students". He applies this representation to the *qud*, instantiates D by the *qud*-domain and gets a semantic representation of the entire sentence: $\forall x[x \in \{y|y \in cs \land french(y)\} \land student(x) \rightarrow talk(x, mary)].$

The determiner "all" is represented as: $\lambda P_1 \lambda P_2 [\forall x [x \in D \land P_1(x) \to P_2(x)]]$. Before this term can be applied to the *qud* of answer 1-2, it has to be transformed into the quantifier $\lambda P_2 [\forall x [x \in D \to P_2(x)]]$. This can be done by application to the trivial predicate $\lambda x [\top]$.² After this type-shifting operation, the quantifier can be applied to the *qud* and the domain-variable can be instantiated, so that the listener derives the intended interpretation of the entire answer: $\forall x [x \in \{y | y \in cs \land student(y)\} \to talk(x, mary)]$.³

A type-shifting operation is needed for the interpretation of the answers 1-3, 2-1 and 3-1, too. "Student" in answer 2-1 is represented as $\lambda x[student(x)]$. The term has to be transformed into a quantifer, before it can be applied to the *qud*. By transforming $\lambda x[student(x)]$ into the existential quantifier $\lambda P[\exists x[x \in D \land student(x) \land P(x)]]$, the listener derives a somewhat weaker interpretation than was intended: $\exists x[x \in$ $\{y|y \in cs \land french(y)\} \land student(x) \land talk(x, mary)]]$. Experiments show that, if an adequate, contextually salient entity s is given, listeners tend to apply an anaphorical interpretation. In this case, the existential quantifier can be strengthend: $\lambda P[\exists x[x \in D \land student(x) \land P(x) \land x = s]].$

The same accentuation can be optimal with respect to different *quds* (cf. example 2.2). If the listener recognizes an entire utterance, then he can (re-)construct the set of questions with respect to which the utterance counts as optimally accentuated. The utterance was in fact optimally accentuated, if the *qud* is an element of this set of background questions (*bq-set*).

Example 2.2

answer The FRENCH student talked to Mary.

qud 1 Which student talked to Mary?

qud 2 What kind of student talked to Mary?

We claim that for the interpretation of not entirely recognized utterances, listeners can use type-shifting operations and apply semantic representations to the *qud*. The entire process

 ${}^{2} \lambda P_{2}[\forall x[x \in D \land \top \to P_{2}(x)]] \Leftrightarrow \lambda P_{2}[\forall x[x \in D \to P_{2}(x)]]$

³ Now it becomes clear why we represent which-questions as whoquestions with a further domain-restriction. If we representes the qud I as $\lambda x[student(x) \wedge talk(x, mary)]$ then answer 1-2 would be interpreted as $\forall x[x \in cs \rightarrow student(x) \wedge talk(x, mary)]$. This would not represent the intended meaning of answer 1-2, since the answer can be true even if some contextually salient person is not a student. Another reason for representing which-questions as described is that this ensures that questions like "Which men are bachelors?", "Which bachelors are men?" and "Who is a man and a bachelor?" evoke different, non-equivalent interpretations. (Cf. [4])

¹ "Yves" is represented as: $\lambda P[yves \in D \land P(yves)]$. The domain-variable D can be instantiated by the domain of the *qud*. If

[&]quot;Yves sleeps" was uttered as an answer to "Which student sleeps?", then its meaning is represented as: $yves \in \{x | x \in cs \land student(x)\} \land$ sleep(yves). According to this representation, the answer asserts that Yves is a student and that Yves sleeps. Intuitively, the answer rather presupposes than asserts that Yves is a student. For simplicity, we do not distinguish between the presuppositional and the assertive parts of an utterance here.

of interpretation can be modelled within a framework of update semantics with questions (cf. [3]).

3. Circumstances of Interpretation

Different circumstances for the interpretation of an utterance can be given: Either the listener recognizes all words of an utterance (case 1), or he recognizes only some words. In the latter case, he may recognize all accentuated words (case 2) or not all of them (case 3). Each of these cases has subcases: If the listener recognizes some accentuated and non-accentuated words, then he might, but not necessarily has to, recognize the difference in accentuation. Finally, either the listener is aware of the *qud*, or he is not.

Case 1-1: The listener correctly recognizes all words of the utterance, he distinguishes between accentuated and non-accentuated words, and he is aware of the *qud*. By using lexicon and grammar, the listener can construct a semantic representation of the utterance and its bq-set. If the *qud* is an element of the bq-set, then the utterance was optimally accentuated. Even if the utterance was not optimally accentuated, it can be understood as an informative answer to the *qud*.

Example 3.1

qud Who is the one who talked to Mary?

answer 1 Yves talked to MARY.

answer 2 SOMEONE talked to Mary.

The first answer of example 3.1 is not optimally accentuated, but, if entirely recognized, an informative answer. Contrary, the second answer of example 3.1 is optimally accentued, but nevertheless it is not an informative answer.

Listeners can draw conversational implicatures from nonoptimally accentuated utterances with respect to the current *qud*. Non-optimal accentuation does not obey Grice's maxim of manner. The violation of a conversational maxim can evoke an implicature. Therefore, non-optimal accentuation with respect to the current *qud* can be conversationally useful. Example 3.2 is from Rooth ([7]).

Example 3.2

context Steve, Paul and Mats took a quiz. Speaker asks Mats.

qud How did it go?

answer 1 Well, I passed.

answer 2 Well, I PASSED.

Mats' first answer is not optimally acccentuated with respect to the *qud*, but to the question who passed the quiz. The listener can take the utterance as an exhaustive answer to that question. He can conclude that Steve and Paul did not pass the quiz. Mats' second answer is neither optimally accentuated with respect to the *qud*. It is optimally accentuated with respect to the question, how it went for Mats. If the listener takes the second answer as exhaustive to this question, he concludes that Mats passed the quiz but did not do better, i.e. he did not ace.⁴

If the listener presupposes that the utterance was optimally accentuated, then he can ignore the non-accentuated constituents and construct a semantic representation from the accentuated constituents by type-shifting and application of the *qud.* By interpreting the utterance that way, the listener might compensate grammatical anomalies, like in example 3.3.

Example 3.3

qud Who talked to Mary?

answer YVES Mary talk.

Interpreting by type-shifting and applying the *qud* might be the favorite strategy to interpret an utterance, because even if the listener correctly recognized all words he need not know that he did. For him, it might be the case that the interpretation without considering the *qud* leads to improper understanding.

Case 1-2: The listener recognizes all words of the utterance, he distinguishes between accentuated and nonaccentuated words, but he is not aware of the qud. He can construct a semantic representation of the utterance and its bqset. If he assumes that the utterance was optimally accentuated, then he assumes that the qud was one of the questions from the bq-set. He can accommodate his knowledge state accordingly. Given the bq-set has only one element, this element must be the qud. Otherwise, the listener can accomodate his knowledge state to a most general bq. If there is only one most general bq, then an exhaustive answer to that bq resolves all other bqs as well. After accomodating the knowledge state, the listener can strengthen the meaning of the utterance by assuming that it is an exhaustive answer to its assumed qud. E.g., he might interpret "YVES talked to Mary" in the sense of "Yves and no other interesting person talked to Mary".

Case 1-3: The listener recognizes all words of the utterance, but he cannot distinguish between accentuated and nonaccentuated words. He is aware of the *qud*. He can construct a semantic interpretation of the utterance. Although he cannot construct a representation of the *bq-set* from the utterance itself, he can decide whether the utterance is an informative answer to the given *qud*, and he can strengthen the meaning of the utterance to an exhaustive answer. He cannot draw implicatures from the accentuation, but only from his knowledge of the *qud*.

Case 1-4: The listener recognizes all words of the utterance, but he cannot distinguish between accentuated and nonaccentuated words. Moreover, he is not aware of the *qud*. He can construct a semantic interpretation of the utterance, but not of its *bq-set*. He has no clue, in which respect the utterance is relevant. He cannot strengthen the utterance's meaning. He cannot draw implicatures from the accentuation.

Case 2-1: The listener does not recognize all words, but all accentuated words of the utterance. He distinguishes between accentuated and non-accentuated words, and he is aware of the *qud*. He can ignore all non-accentuated words and construct a semantic representation of the complete utterance by means of type-shifting and application of the *qud*. The listener can construct underspecifying representations of *bqs*. To that end, he replaces words that are syntactically required but have not been recognized, by variables of the appropriate types. If the underspecifying representation of one *bq* und the *qud* can be unified (cf. [6]) – i.e. if one *bq* can be specified so that it becomes identical to the *qud* –, then the listener has good evidence that the utterance was optimally accentuated and that the derived interpretation is correct. The listener can strengthen the derived meaning to an exhaustive answer.

Case 2-2: The listener does not recognize all words, but all accentuated words of the utterance. He distinguishes between accentuated and non-accentuated words, but he is not aware of the *qud*. The listener can construct underspecifying representations of the utterance meaning and the *bqs*. If he does not have a

 $^{^4}$ Rooth ([7]) calls these implicatures *scalar implicatures* and describes them as effects of focussing. In this paper, we do not rely on a notion of *focus* distinct from accent. We describe the implicatures as effects of mere accentuation.

further clue on how to specify these representations, he cannot understand the entire utterance.

Example 3.4

qud Who talked to Mary?

answer 1 YVES talked to noise

answer 2 YVES noise Mary.

continuation PETER did NOT talk to Mary.

From the recognized parts of the first answer of example 3.4, the listener can derive that it is under discussion who talked to someone or noone, and that Yves was the one who talked. The second argument of "talked to" has to be further specified. If the listener assumes that Mary is in the center of attention, then he might specify the derived utterance-meaning to that of "Yves talked to Mary". Otherwise, he might ask for clarification: "Yves talked to WHOM?" From the recognized parts of the second answer, the listener can derive that Yves stands in some relationship to Mary and that it was under discussion who stands in this relationship to Mary. One bq of the continuation is who talked to Mary. If the listener fully recognizes the continuation after the second answer, he might assume that both sentences answer the same question. He therefore might specify the meaning of the second answer to that of "Yves talked to Mary". He does not have to ask the speaker for clarification.

Case 2-3: The listener does not recognize all words, but all accentuated words of the utterance. He does not distinguish between accentuated and non-accentuated words, but he is aware of the *qud*.

Example 3.5

qud *Who talked to Mary?*

answer 1 Yves noise

answer 2 Yves talked to noise

answer 3 Yves noise Mary.

If the listener recognizes exactly the accentuated words, then he can easily construct a semantic representation by type-shifting and the application of the *qud* (example 3.5, answer 1). If that does not work, because the listener recognized more than the accentuated words, then he has to ignore some words and try to construct a semantic interpretation again. Basically, interpretation works as in case 2-1, with the difference that the listener has no clue which words can be ignored. He has to interpret by trial and error (example 3.5, answer 2). The selection of words can be guided by plausibility assumptions. So it seems promising to identify the recognized "Mary" in answer 3 (example 3.5) with the occurrence of "Mary" in the *qud*.

Case 2-4: The listener does not recognize all words, but all accentuated words of the utterance. He does not distinguish between accentuated and non-accentuated words, and he is not aware of the *qud*. Misunderstanding is probable. The best strategy for the listener might be to assume that all recognized words were accentuated and to interpret the utterance along the lines of case 2-2.

Case 3: The listener does not recognize all accentuated words of the utterance. Lack of understanding is highly probable.

Example 3.6

qud Who talked to Mary?

(intended answer: The french student talked to Mary.)

answer 1 noise talked noise

answer 2 noise Mary noise

answer 3 noise student noise

The first answer of example 3.6 cannot be properly interpreted. The listener has to ask the speaker to repeat. The listener can construct a semantic representation from the second answer, if he is aware of the *qud*. But this representation would stand for the meaning of "Mary talked to Mary", which is not the intended meaning. Again, misunderstanding is probable. From the third answer and the *qud*, the listener can construct a weak semantic representation ("Some student talked to Mary"). This representation might be satisfying, although it does not fully represent the intended meaning.

4. Conclusions

The less the listener is aware of the context of an utterance and the fewer words the listener recognizes, the more difficult is the interpretation and the more probable is misunderstanding. Anyway, interpretation is robust. Even with incomplete awareness of the context and incomplete recognition of an utterance, the listener might be able to correctly interpret the entire utterance.

Clear articulation involves accentuating those words which the listener must recognize in order to understand the entire utterance. Optimal accentuation is one means of clear articulation. Speakers do not always accentuate optimally. Nevertheless, it seems to be a good and common strategy of interpretation to follow the *principle of charity* and assume that a given utterance was optimally accentuated.

In some cases, accents seem to influence the semantic meaning of utterances. Such phenomena are commonly described as *associations with focus* ([7]). The plausibility of our hypothesis rises, if associations with fous can be explained as epiphenomena of optimal accentuation. Within an adequate framework of formal pragmatics such an explanation seems feasible.

5. References

- Ginzburg, J., 1996. Dynamics and the Semantics of Dialogue. In *Logic, Language and Computation, Volume 1*, Seligman, J. and Westerståhl, D. (eds.). Stanford: CSLI, 221-237.
- [2] Grice, P., 1967. Logic and Conversation. In *Studies in the Way of Words*, Grice, P., 1989. Cambridge/ Mass: HUP, 1-143.
- [3] Groenendijk, J., 1999. The Logic of Interrogation, Classical Version. *Ninth Conference on Semantics and Linguistic Theory*, Santa Cruz.
- [4] Higginbotham, J., 1996. The Semantics of Questions. In *The Handbook of Contemporary Semantic Theory*, Lappin, S. (ed.). Oxford: Blackwell, 361-383.
- [5] van Kuppevelt, J., 1995. Discourse Structure, topicality and questioning. *Journal of Linguistics*, 31, 109-147.
- [6] Pullman, S., 1997. Higher Order Unification and the Interpretation of Focus. *Linguistics and Philosophy*, 20, 73-115.
- [7] Rooth, M., 1992. A theory of focus interpretation. *Natural Language Semantics*, 1, 75-116.