

Scripted Simulated Dialogue: a new elicitation paradigm

Glasbergen-Plas, A.; Gryllia, Stella; Pablos Robles, Leticia; Doetjes, Jenny

DOI

[10.21437/SpeechProsody.2022-139](https://doi.org/10.21437/SpeechProsody.2022-139)

Publication date

2022

Document Version

Final published version

Citation (APA)

Glasbergen-Plas, A., Gryllia, S., Pablos Robles, L., & Doetjes, J. (2022). *Scripted Simulated Dialogue: a new elicitation paradigm*. 683-687. Paper presented at Speech Prosody 2022, Lisbon, Portugal.
<https://doi.org/10.21437/SpeechProsody.2022-139>

Important note

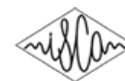
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



Scripted Simulated Dialogue: a new elicitation paradigm

Aliza Glasbergen-Plas¹, Stella Gryllia², Leticia Pablos Robles³, Jenny Doetjes⁴

¹Delft University of Technology, Netherlands

^{2,3,4}Leiden University, Netherlands

a.glasbergen-plas@tudelft.nl, gryllia.stella@gmail.com

Abstract

Two methods are commonly used to elicit production data for prosody research. The first, in which participants read out a series of written sentences, gives good control over what data are elicited. The second, in which participants perform a task designed to elicit the speech of interest (e.g., a Referential Communication Task), is suitable for studying speech in context. However, certain research topics require the combination of these qualities. We developed an elicitation paradigm, Scripted Simulated Dialogue, that (a) gives precise control over the data that are elicited and (b) is suitable for studying speech in context. In addition, it allows the researcher to control or manipulate the preceding discourse, whereas a Referential Communication Task provides discourse that may be analysed afterwards. The paradigm simulates a series of short dialogues, in which the participant reads her text from a screen and the ‘interlocutor’ is a recorded voice. The participants are not made aware of which speech turn in the dialogue contains the target sentence. We illustrate how Scripted Simulated Dialogue may be used to manipulate the context and make the E-Prime script available to other researchers.

Index Terms: Scripted Simulated Dialogue, elicitation paradigm, context, E-Prime script

1. Introduction

Two methods are commonly used to elicit production data for prosody research: scripting and communication tasks like a Referential Communication Task [1] or a Discourse Completion Task [2]. While these methods yield good results, certain research topics require a third option that combines precise control over the elicited data (as with scripting) with the possibility to study speech in context (as with communication tasks) (cf., [3]). For research topics requiring this combination, we introduce Scripted Simulated Dialogue, for which we make available an E-Prime Script.

We first lay out how Scripted Simulated Dialogue complements the current common elicitation methods in Section 2. Section 3 describes the properties of the new paradigm and Section 4 illustrates how it can be used to manipulate the context preceding a target utterance. Finally, the paper describes the paradigm’s drawbacks and benefits in Section 5 and the E-Prime Script in Section 6.

2. Common elicitation methods

We start by explaining how Scripted Simulated Dialogue adds to the commonly used elicitation methods: scripting and communication tasks.

2.1. Scripting

With scripting, the experimenter presents participants with a series of pre-scripted sentences to read out loud. The sentences are presented in isolation (e.g., [4], [5]), or after a short fragment of context (e.g., [6], [7]). An example of scripting using a short fragment of context is Dohen & Lœvenbruck’s [6] study of contrastive focus in French. One of the target sentences in this study was *Les loups suivaient MARILOU* ‘The wolves followed MARILOU’, with intended contrastive focus on the name. To elicit this target, a participant would first see the sentence *Les loups suivaient Marilou*. Next, (s)he would hear a prompt: *Les loups suivaient Aurélie?* ‘Did the wolves follow Aurélie?’ The prompt was varied to create different focus structures in different conditions. While such a prompt is a form of context, the preceding discourse other than the prompt is not specified. Scripting gives very good control over the segmental composition of the elicited data and easily allows comparisons across conditions.

2.2. Communication tasks

Common communication tasks include Referential Communication Tasks (RCT) and Discourse Completion Tasks. A Referential Communication Task [1], like Brown et al.’s [8] map task, involves two participants or a participant and a confederate performing some task together. The task is designed to elicit a conversation that contains the speech of interest. An example of an RCT is Ito & Speer’s [9] ‘tree decoration task’. It was used to study the relation between tonal patterns and the information status of words in English and Japanese. The ‘tree decoration task’ has a participant instructing another person about how to decorate a Christmas tree. The participant points out objects to hang in the tree, such as a small orange drum, which elicits an utterance containing the target *orange drum*.

In a Discourse Completion Task [2], participants are presented with a scenario that contains background information about an event, as well as information on the social distance between the interlocutors. Participants are then asked to respond by completing a turn of dialogue.

The strength of communication tasks is that they elicit an actual dialogue in the laboratory. Participants are assumed to focus on the task at hand or the event, rather than on the form of their speech. Moreover, this method allows the researcher to study speech in context.

2.3. Scripted Simulated Dialogue’s contribution

Scripted Simulated Dialogue combines elements of both methods described above. It uses scripting, but also simulates

(to some extent) a conversation. Its contribution is therefore that it combines control over the elicited data with the presence of a discourse context. In addition, it allows for control or manipulation of the preceding discourse, in contrast to a Referential Communication Task, which provides uncontrolled discourse that may be analysed afterwards.

3. Scripted Simulated Dialogue

Scripted Simulated Dialogue simulates a series of short dialogues, in which the participant's interlocutor is a recorded voice. The participant's speech turns are scripted: s/he reads them from a computer screen. An example of a dialogue (in French) is presented in Figure 1. Speaker A represents the participant and Speaker B the 'interlocutor'. The target sentence is underlined here, but it would not be in an actual experiment.

[Conversational setting] Tu discutes avec Ernestine, ta femme. Elle part quelques jours en voyage d'affaires et rentrera mercredi, juste à temps pour ton anniversaire. Tu lui dis :

- A *Bon voyage ma chérie. Tu as bien ton passeport ?*
 B *Oui merci. Ah voilà mon taxi.*
 A *Tu m'envoies un texto quand tu es arrivée à Londres ?*
 B *Oui, oui bien sûr. A mercredi ;
 pour ta dernière soirée de trentenaire !*
 A *Moque-toi ; dans six mois c'est ton tour.
 D'ailleurs tu ne m'as pas dit.
Tu as réservé quel resto pour jeudi soir ?*
 B *Surprise...*

[Conversational setting] You're talking to Ernestine, your wife. She's going on a business trip for a few days and will be back on Wednesday, just in time for your birthday. You say:

- A *Have a good trip love. Have you got your passport?*
 B *Yes thanks. Oh that's my taxi.*
 A *Will you send me a text
 when you've arrived in London?*
 B *Yes, sure, I will. On Wednesday;
 on your last evening in your thirties!*
 A *Careful with the teasing; in six months it's your turn.
 By the way, you didn't tell me.
Which restaurant did you book for Thursday evening?*
 B *Surprise...*

Figure 1: Example of a Scripted Simulated Dialogue in French with English translation underneath [12, p. 581].

As in Figure 1, every dialogue is preceded by a description of the conversational setting, which states who the interlocutors are and where the conversation takes place. The context manipulation thus has two elements: this description and the preceding speech turns. Every dialogue has one target sentence or filler embedded in it, always at the same position in the dialogue. The participant does not know in which speech turn. As the target is positioned towards the end of the dialogue, the preceding discourse can be used to manipulate a particular reading of the sentence.

After each dialogue, the participant receives a question about the information supplied by the recorded 'interlocutor', as in (1).

- (1) Ernestine rentre de voyage...
 1. mercredi
 2. samedi
 3. vendredi

Ernestine is coming back from her trip on...

1. Wednesday
 2. Saturday
 3. Friday

The purpose of this is to direct the participants' attention to the content of the dialogue, rather than the form of the utterances.

The procedure during an experiment can be as follows. A participant presses a key once s/he is ready to start. This prompts the recording of the first conversational setting to be played through headphones, while the screen is blank. Then the participant's first speech turn appears on the screen. The participant utters his/her speech turn, after which s/he presses a key for the 'interlocutor's' speech turn to start playing through the headphones while the screen is blank again. Then the participant's next speech turn appears on the screen. This process is repeated until the participant has produced the utterance that forms the target sentence in his/her third speech turn (or a filler) and received a reaction in the 'interlocutor's' third and last speech turn. After the last speech turn of every dialogue, a multiple-choice sentence completion task as in (1) above appears on the screen, asking about information supplied by the 'interlocutor'. The participant answers the question by pressing 1, 2 or 3. Feedback on the answer may appear on the screen. The participant then presses a key to move on to the next trial. The E-Prime [10] script discussed in Section 5 produces this sequence. We recommend using a male voice for the conversational setting and a female voice for the 'interlocutor's' speech turn, or vice versa.

4. Illustration: two context manipulations

We will now illustrate how the paradigm may be used to manipulate the discourse context preceding a target utterance. To this end, we summarise a production experiment that targeted the prosody of French *wh*-in-situ questions elicited in different types of context [11], [12]. The study manipulated (a) the focus structure of the target sentences and (b) the distinction between echo and information seeking questions. We first introduce the topic of the study (Section 4.1) and the experimental stimuli (Section 4.2). Then, we present the ways in which Scripted Simulated Dialogue was used to manipulate the context (Section 4.3).

4.1. French *wh*-in-situ questions and context

French has multiple ways to form a *wh*-question. In (2a), the *wh*-phrase *qui* 'who' has been moved to the left periphery of the sentence, while in (2b), it occupies the same position as in the corresponding declarative in (2c): it is left 'in-situ'.

- (2) a. Qui Jean a-t-il invité ?
 who Jean has-T-he invited
 'Who did Jean invite?'
 b. Jean a invité qui ?
 Jean has invited who
 'Who did Jean invite?'
 c. Jean a invité Pierre.
 Jean has invited Pierre
 'Jean invited Pierre.'

Wh-in-situ questions as in (2b) are not part of the prescriptive grammar, but are frequently attested, often but not exclusively in an informal register [13], [14].

Questions of this form also occur as echo questions, ‘echoing’ the previous utterance (3). In (3), part of speaker A’s utterance was not clearly audible, prompting speaker B to ask for a repetition.

- (3) A: Jean a invité #####[noise].
 Jean has invited #####[noise]
 ‘Jean invited #####[noise].’
 B: Jean a invité qui ?
 Jean has invited who
 ‘Jean invited who? (I did not hear you.)’

Whereas some authors suppose that the *wh*-phrase always equals the focus in *wh*-questions (e.g., [15], [16]), others suppose that the focus structure of a *wh*-question, at least in some languages, depends on the preceding context (e.g., [17, 18]). Glasbergen-Plas [11] and Glasbergen-Plas et al. [12] follow this latter point of view (see [11, Ch. 3] for discussion).

The production experiment investigated whether (a) the focus structure and (b) the distinction between echo and information seeking questions is reflected in the prosody of French *wh*-in-situ questions. The context preceding a question was manipulated to create the conditions in (4). Echo questions always have a narrow focus on the *wh*-word (see [11, Ch. 3]). Therefore, Condition C represents an information seeking question with the same focus structure as an echo question.

- (4) CONDITIONS
 A. Echo question (expressing auditory failure)
 B. Information seeking question with broad focus
 C. Information seeking question with narrow focus

4.2. Stimuli

The target stimuli had a form as in (5), in which the first translation represents an information seeking question and the second an echo question.

- (5) Tu as réservé quel resto pour jeudi soir ?
 you have booked which restaurant for Thursday evening?
 IS: ‘Which restaurant did you book for Thursday evening?’
 E: ‘You booked which restaurant for Thursday evening?’

The study employed twelve target stimuli as in (5). Each of these was presented in the three conditions in (4), yielding a total of thirty-six target utterances. They were intermingled with thirty-six fillers.

Each stimulus or filler was embedded in a dialogue as in Figure 1 above, with three speech turns for the participant and three for the recorded ‘interlocutor’. The stimulus or filler was always part of the participant’s last speech turn, with the ‘interlocutor’s’ last speech turn following it.

4.3. Context manipulations

We will now describe how the preceding context was used in this study to trigger an echo question expressing auditory failure (Condition A), an information seeking question with broad focus (Condition B) or an information seeking question with narrow focus on the *wh*-word (Condition C).

Figure 2 presents an example of a dialogue used in Condition A (echo question). Pink noise covers the word

Monette, represented as strikethrough text. This causes a need to ask for repetition. An episode of pink noise was also present in all other contexts (pertaining to the other conditions and the fillers), but in a position where it would not hinder the conversation, for instance on the final syllable of a long word.

[Conversational setting] Tu es directeur d’une petite école primaire. La semaine prochaine, c’est la rentrée des élèves. Mais, demain, mercredi, c’est la pré-rentrée pour les maîtres et maîtresses. Tu es à l’école avec Axelle, ta secrétaire, pour organiser les dernières petites choses. Tu dis :

- A *Et c’est reparti pour un an !*
 B *Oui et avec deux classes et deux nouvelles maîtresses de plus.*
 A *C’est bien qu’on ait prévu ce petit dîner pour faire plus ample connaissance.*
 B *Oui, d’ailleurs je voulais te dire, pour qu’on soit au calme pour parler, j’ai réservé le resto « chez ~~Monette~~ » pour jeudi soir.*
 A *Tu as réservé quel resto pour jeudi soir ?*
 B *Chez Monette, dans la petite salle du fond, on devrait être tranquilles.*

[Conversational setting] You are the principal of a small primary school. Next week, it’s the start of the new school year. But tomorrow, Wednesday, is the first day for the teachers. You are at the school together with Axelle, your secretary, to organize the last things. You say:

- A *So we start again!*
 B *Yes, and with two new classes and two new teachers.*
 A *It was a good idea to have this small dinner party to get to know each other.*
 B *Yes, by the way, I wanted to tell you. In order to have a quiet place to talk, I booked the restaurant ~~Chez Monette~~ for Thursday evening.’*
 A *You booked which restaurant for Thursday evening?*
 B *Chez Monette. They have a back room that’s usually quiet.*

Figure 2: Scripted Simulated Dialogue used in Condition A (echo question) with English translation underneath [12, p. 580].

To elicit information seeking *wh*-in-situ questions with broad focus (Condition B), the study used dialogues as in Figure 1 on the previous page. Although the target sentence was preceded by context in this condition, it provided little information about the content of the question. Whereas the context was consistent with the existential presupposition or implicature of *wh*-questions (i.e., the speaker expected there to be an answer), no part of the content of the question would be mentioned in the preceding context. Consequently, the *wh*-in-situ question formed a rather sudden departure from the topic of the preceding conversation. To keep the discourse natural, the context signalled this change in topic, for instance by a ‘topic change marker’ [19], like *d’ailleurs tu ne m’as pas dit* ‘by the way, you didn’t tell me’ (see Figure 1).

In Condition C, the context was designed to force a reading as an information seeking question with narrow focus on the *wh*-word: the same information structure as an echo question. To this end, the context would mention all elements of the

content of the question except the *wh*-word, such as ‘booking a restaurant for Thursday evening’ in Figure 3.

[Conversational setting] Tu es violoniste dans un orchestre amateur. Tu es en séance de répétition. Pendant que les flûtistes répètent un passage délicat, tu parles avec ta voisine Eléonore. Tu lui dis :	
A	<i>Tu pars en déplacement cette semaine ?</i>
B	<i>Non pour une fois, je suis là toute la semaine. Ça tombe bien, c'est la remise de diplôme de ma fille jeudi. Du coup, nous allons en famille au restaurant.</i>
A	<i>C'est marrant, Fleur m'a raconté la même chose.</i>
B	<i>Oui, elle m'a dit qu'elle a réservé au Pavillon pour jeudi soir.</i>
A	<i>Et toi, <u>tu as réservé quel resto pour jeudi soir ?</u></i>
B	<i>Le Bord du Lac.</i>

[Conversational setting] You play the violin in an amateur orchestra. During a rehearsal, while the flutists are practicing a particularly difficult passage, you talk to Eléonore, who is sitting next to you. You say:	
A	<i>Are you going on a trip this week?</i>
B	<i>No, just this once I'm going to be here all week. Good timing: it's my daughter's graduation ceremony on Thursday, so we're going out for a family dinner.</i>
A	<i>Oh that's funny, Fleur said just the same thing.</i>
B	<i>Yes, she told me she'd booked the restaurant Pavillon for Thursday evening.</i>
A	<i>And you, <u>which restaurant did you book for Thursday evening?</u></i>
B	<i>Le Bord du Lac.</i>

Figure 3: *Scripted Simulated Dialogue used in Condition C (narrow focus) with English translation underneath [12, p. 581].*

In order to create this type of context while keeping the flow of the discourse natural, the study used *wh*-in-situ questions with a contrastive topic, as in [20, p. 100]. Subject pronouns in French are clitics and cannot be contrastively stressed [21]. To express contrastive topichood, French uses a left dislocated ‘strong’ pronoun, which is coreferential with a clitic [22, pp. 115–116]. The study used *et toi* ‘and you’, which was taken up by the resumptive clitic *tu* ‘you’ in the clause proper. Consequently, the part of the sentence following the contrastive topic *et toi* ‘and you’ was string-identical to the target stimuli used in Conditions A and B.

5. Drawbacks and benefits

The Scripted Simulated Dialogue paradigm has two clear drawbacks. First, it does not involve actual conversation in the lab as with communication tasks, since the dialogues are pre-scripted and participants read them from a screen. In actual dialogue, speakers may know in advance what they are going to say, which could affect the prosody. Second, the pre-recorded speech turns of the ‘interlocutor’ may prime the participant.¹

¹ We thank an anonymous reviewer for pointing this out.

However, Scripted Simulated Dialogue has several benefits:

- It combines the experimental control that comes with scripting with the possibility to include context.
- It allows for manipulation of the context. This is a clear advantage over a Referential Communication Task, which provides uncontrolled discourse that can only be analysed afterwards.
- We suggest that it has better ecological validity than mere scripting. For instance, the fact that the target sentence is embedded in a larger dialogue reduces the chance that the participant becomes aware of the topic of study.

Marandin [3] points out that the way in which an utterance is read can be influenced by a speaker’s interpretation or imagination of the discourse situation and his/her motivation regarding what and how much to answer. Scripted Simulated Dialogue forms a way to control for this.

The paradigm is particularly useful to elicit targets that require context yet cannot be elicited in a communication task. These are targets of which the meaning or a similar meaning can be expressed in multiple ways. Ito & Speer’s ‘tree decoration task’ (see Section 2.2) successfully elicits NPs. However, it is almost impossible to set up a communication task that predictably elicits the target utterances in Figures 1–3. This is because the meaning expressed by these targets can also be expressed in other ways, such as one of the other types of *wh*-question in French (cf., (2a)). In addition, participants could express an alternative meaning that also fits the context, such as ‘Are you going out in the city centre?’ or ‘You like Le Bord du Lac, don’t you, are you going there for the dinner?’

6. E-Prime script

We make available an E-Prime script for Scripted Simulated Dialogue (programmed in E-Prime 2.0) that may be adjusted to the needs of the researcher. The script in the zipped folder *ScriptedSimulatedDialogue*² contains three blocks of 24 dialogues (12 items and 12 fillers), preceded by three practice trials and instruction screens. The text of the instruction screens, the speech turns of the participants and the sentence completion tasks can be typed in. Links to sound files can be inserted. We recommend using authentic sounding dialogues that fit the required register.

7. Acknowledgments

The paradigm has been developed as part of the project Understanding Questions (360-70-480), funded by the Dutch Research Council (NWO). We wish to thank Lisa Cheng and Yang Yang for fruitful discussions on the research paradigm and Yannick Gloanec, Marion Bracq and in particular Sylvie Cuchet for their help in creating the French dialogues.

8. References

- G. Yule, *Referential Communication Tasks*. Mahwah, NJ: Lawrence Erlbaum Associates, 1997.
- M. del Mar Vanrell, I. Feldhausen, and L. Astruc, “The Discourse Completion Task in Romance prosody research: Status quo and outlook,” in *Methods in prosody: A Romance language perspective*, I. Feldhausen, J. Fließbach, and M. d. M. Vanrell, Eds. Berlin: Language Science Press, 2018, pp. 191–227.

² We included both the .ebs2 and the .es2 file. A license for E-Prime is required.

- [3] J.-M. Marandin, "Affordance and ability: How do participants replicate linguistic choices in the lab?," *Belgian journal of linguistics*, vol. 25, no. 1, pp. 30-50, 2011.
- [4] S. Gryllia, L. L.-S. Cheng, and J. Doetjes, "On the intonation of French *wh*-in-situ questions: What happens before the *wh*-word is reached?," presented at the Speech Prosody 2016, Boston, MA, 2016.
- [5] M. d'Imperio and A. Michelas, "Pitch scaling and the internal structuring of the Intonation Phrase in French," *Phonology*, vol. 31, no. 1, pp. 95-122, 2014.
- [6] M. Dohen and H. Lævenbruck, "Pre-focal rephrasing, focal enhancement and postfocal deaccentuation in French," presented at the Proceedings of the 8th International Conference on Spoken Language Processing (Interspeech 2004), Jeju Island, Korea, 2004.
- [7] C. Beyssade, B. Hemforth, J.-M. Marandin, and C. Portes, "Prosodic Realizations of Information Focus in French," in *Explicit and Implicit Prosody in Sentence Processing: Studies in Honor of Janet Dean Fodor*, L. Frazier and E. Gibson, Eds. New York City: Springer, 2015, pp. 39-61.
- [8] G. Brown, A. Anderson, R. Shillcock, and G. Yule, *Teaching talk: Strategies for production and assessment*. Cambridge: Cambridge University Press, 1984.
- [9] K. Ito and S. R. Speer, "Using interactive tasks to elicit natural dialogue," in *Methods in empirical prosody research*, S. Sudhoff *et al.*, Eds. Berlin: Walter de Gruyter, 2006, pp. 229-257.
- [10] Psychology Software Tools Inc. (2012). [*E-Prime 2.0*], 4 January 2016, <http://www.pstnet.com>. [Online]. Available.
- [11] A. Glasbergen-Plas, *Questions in context: the case of French wh-in-situ*. Amsterdam: LOT publications, 2021.
- [12] A. Glasbergen-Plas, S. Gryllia, and J. Doetjes, "The prosody of French *wh*-in-situ questions: echo vs. non-echo," *Journal of Linguistics*, vol. 57, no. 3, pp. 569-603, 2021.
- [13] V. Quillard, "Interroger en français parlé: études syntaxique, pragmatique et sociolinguistique," Ph.D. dissertation, Université de Tours, 2000.
- [14] L. L. Myers, "Wh-interrogatives in spoken French: A corpus-based analysis of their form and function," Ph.D. dissertation, The University of Texas at Austin, 2007.
- [15] P. W. Culicover and M. Rochemont, "Stress and focus in English," *Language*, vol. 59, no. 1, pp. 123-165, 1983.
- [16] K. Lambrecht and L. A. Michaelis, "Sentence accent in information questions: Default and projection," *Linguistics and Philosophy*, vol. 21, no. 5, pp. 477-544, 1998.
- [17] J. Jacobs, "Implikaturen und 'alte Information' in *w*-Fragen," in *Fragesätze Und Fragen*, M. Reis and I. Rosengren, Eds. (Linguistische Arbeiten, no. 257). Tübingen: Niemeyer, 1991, pp. 201-221.
- [18] R. Eckardt, "Inherent focus on *wh*-phrases," in *Sinn und Bedeutung 11*, E. Puig-Waldmueller, Ed. Barcelona: Universitat Pompeu Fabra, 2007, pp. 209-228.
- [19] B. Fraser, "What are discourse markers?," *Journal of Pragmatics*, vol. 31, no. 7, pp. 931-952, 1999.
- [20] E. Engdahl, "Information packaging in questions," in *Empirical Issues in Formal Syntax and Semantics*, vol. 6, O. Bonami and P. C. Hofherr, Eds. no. 1). Paris: CSSP, 2006, pp. 93-111.
- [21] R. S. Kayne, *French syntax: The transformational cycle*. Cambridge, MA: MIT Press, 1975.
- [22] K. Lambrecht, *Information Structure and Sentence Form*. Cambridge: Cambridge University Press, 1994.