'No No No No No No No': Multiple Repetition in Dialogue

Zulipiye Yusupujiang[1]   **Paola Herreno Castaneda**[2]   **Mathilde Dargnat**[2]   **Jonathan Ginzburg**[1]

[1]CNRS, Université Paris-Cité   [2]CNRS, Université de Lorraine
Laboratoire de Linguistique Formelle   ATILF

{zulipiye.yusupujiang,yonatan.ginzburg}@u-paris.fr
{paola.herreno-castaneda,mathilde.dargnat}@univ-lorraine.fr

Abstract

Disallowing non-modifier words and phrases to occur in duplicate copies (*repetition sequences* (RS)) is a basic principle of formal grammar. The existence of such sequences is pervasive in conversation, with a wide variety of meanings, across diverse languages (data is provided in this paper from English, French, German, Hebrew, Greek, and Chinese). We relate such uses to its even more pervasive occurrence in head gestures, across all three axes (nodding, shaking, tilting). Although the existence of the phenomenon has been noted in previous literature, no systematic taxonomy nor formal treatment has been offered. We offer both a taxonomy with high coverage, but also ground the meanings RS can bear in a dialogical grammar. We also offer a characterization of the words that are particularly susceptible to several classes of RS uses (above and beyond self-repair and emphasis, which seem to be, to a first approximation, unrestricted).

1 Introduction

Formal grammars recognize the possibility of sequences of repeated words for modifiers like adjectives and adverbs via categories of the form $X\backslash X$. On the basis of this, one can capture the grammaticality of examples like (1), the meaning is a more complex story:

(1) a. Peter: but if you went through the Independent and you counted the erm number of different words the total vocabulary, I think you'd find it's **many many many many many many many** times bigger than the . . . (unfinished) [BNC, J40, L190]

   b. Anon3: **Very very very** briefly Chairman, erm first of all you can see in paragraph two point three . . . [BNC, J42, L2]

On the other hand, such grammars will typically rule out cases like (2):[1]

(2) a. Bo Bo left.

   b. Millie likes likes Bo.

However, in spoken language, **repetition sequences (RSs)** of this kind occur across a wide range of categories, words, phrases, and clauses:

(3) a. Grace : Different beat, different beat innit? Anon 3: and it goes ⟨ singing ⟩ 'for ever and ever' **Yeah yeah yeah we know we know** [BNC, KPE, L23 – L25]

   b. dennis mccarthy: **Yes yes.** You've never had to eat fire to get . . . rob tomlinson: **No no no.** [BNC KM2, L821 – L823]

   c. (attested) A: Funny funny because he doesn't really call for a ceasefire or anything B: He does ! 'Immediate stop to fighting'. A: **true true** I hastily read.

   d. Emmy: And I think, yes we will do all that but not quite to the extreme and we should ⟨ unclear ⟩ try now and sell more during the year and really see how the flow goes. Emmy: So instead of **pushing pushing pushing**, let's just see how much comes more or less naturally through advertising. ⟨ unclear ⟩ [BNC, J9P, L202 – L203]

   e. Chris: Hi Susan how are you? Susan: Hi Chris. Fine thank you. Chris: **Good good good.** Geared up for the hols? [BNC, KBK, L1293 – L1297]

---

[1]An anonymous reviewer is skeptical about this claim, suggesting that formal grammars in general do not address the kind of repetition, but do not rule it out. We think they intend to rule such cases out by not providing a 'legal derivation' thereof in their rule/principle-set.

Indeed strikingly some cases of RS are clearly not equivalent to the single uses:

(4) (In a courtroom) A: Were you present at the meeting on 19 November? B: # No, no no / #Yes, yes, yes.

Repetition can also be infelicitous, e.g., RS of backchannels:[2]

(5) A: She told me to leave B: mmh / #mmh mmh A: and then she insulted me.

Floricic and Mignon (2007), in their study of the reduplication of *no* (in French and Italian), point out the different syntactic and semantic constraints of this phenomenon. They specify two main cases in which the utilisation of the form reduplicated seems impossible: in a polar question (with the expression "or no?") as in (6), and in a subordinate after the pronoun "that" as in (7).

(6) a. Ouais et vous êtes parties vous à Disney-land **ou non** ? [CODIM-MPF] *(Yeah and did you went to Disneyland **or no**?)*

   b. ? Ouais et vous êtes parties vous à Disney-land **ou non non ?** *(Yeah and did you went to Disneyland **or no no**?)*

(7) a. Je trouve **que non** personnellement j'aime mieux [CODIM-ESLO] *(I think **that no**, personally I like it better)*

   b. ? Je trouve **que non non** personnellement j'aime mieux *(I think **that no no**, personally I like it better)*

Also, they specify that the reduplication form cannot appear in the final position when this structure opens an answer, as it is shown in (8).

(8) a. A: en ça n'est pas votre mari qui qui lui écrit ? – B: **non non non** c'est toujours moi [CODIM-ESLO] *(A: So it's not your husband who's writing to him? – B: **No no no**, it's always me.)*

   b. ? A: en ça n'est pas votre mari qui qui lui écrit ? – B: c'est toujours moi **non non non** *(A: So it's not your husband who's writing to him? – B: it's always me, **no no no**)*

Repetition has been widely studied, as we discuss in section 2, and, as we will exemplify, the phenomena at issue seem universal (apart from English, we provide data from Chinese, German, Greek, and Hebrew). However, there has been no careful taxonomy of RSs, nor any formal treatment. RSs pose problems both for formal grammars, which are set up to disallow them, but also for dynamic semantic treatments, where repetition at best has a vacuous effect (see e.g., (Kamp et al., 2010).

We start in section 2 with a brief literature review, illustrating that RSs are not formally described in the literature. We then propose a taxonomy in section 3. On the basis of this, we perform a corpus study of the distribution of RS and a classification of its possible distinct uses. We undertake this in section 4. Drawing on its results, we offer a formal grammatical account in section 5, after which we discuss a possible explanation for the words used most frequently for certain classes of RS. Section 7 provides conclusions and future work.

2 Literature Review

Repetition is a frequent phenomenon in spoken discourse, and it has been widely studied within pragmatic and conversational frameworks. Studies found that repetition not only occurs to signal disfluency or redundancy, but it also performs other important functions in discourse. Among earlier studies, scholars observed that repetition reinforces previously stated information (Kernan, 1977) and improves the precision in speech (Erickson, 1984). Bublitz (1989) suggests that repetition helps maintain the continuous and smooth flow of speech, signals the speaker's stance toward what is being said, and also assists both speaker and listener in comprehension. Furthermore, Norrick (1987) offers a nuanced classification of self-repetition, dividing its functions into four categories—**Semantically-based repetition**, **Production-based repetition**, **Comprehension-based repetition**, **Interaction-based repetition**. However, his categories are quite a bit broader than RS, including coordination, reformulation, and mere repetition of words across turns. A taxonomy similar in its nature is that of Rabab'ah and AbuSeileek (2012).

---

[2]An anonymous reviewer expresses doubt about this judgement, suggesting it is felicitous as long as they are intonationally separate. This is distinct from the intonation in typical RS, as in (3), but clearly this calls for a careful experimental study.

Tannen (1989) identifies several functions of repetition in terms of establishing coherence and interpersonal involvement. She suggests that repetition can serve as a mechanism for participatory and ratifying listenership, and also has functions such as stalling, savouring humour, expanding on prior content, and encouraging involvement in conversation. In a study on Mandarin conversational data, Hsieh (2011) also illustrates the role of self-repetition in expressing emphasis, persuasion, and clarification, and points out that it can be used to "double up the illocutionary force, i.e., to do emphasis or to do persuasion, by means of repeating the linguistic form."

(9) a. A: Na ni zenme gen pengyou jieshi, ni zhege haizi shi nali laide ne? (*Then, how do you explain it to your friends where this child came from?*)
B: Um, **wo cong lai bu jieshi. Wo conglai bu jieshi. Wo conglai bu jieshi.** (*Um, I never explain it. I never explain it. I never explain it.*) [Mandarin example (3) from Hsieh (2011) on pages 154-155.]

Dostie (2007, 2011) differentiates repetition from reduplication. Repetition implies a change in the semantic traits between the first and the second element, using each **discourse marker (DM)** to accomplish different actions. As she explains in example (10), the first *là* is used as a spatial deixis while the second one is used as a DM of (discourse unit) segmentation.

(10) Vous allez vous asseoir **là là**. [Dostie, 2007] (*You are going to sit **there there***)

This also agrees with the conversational approach of Stivers (2004), in which the repetition does not share the same prosodic contour. On the contrary, what Dostie (2007) calls pragmatic reduplication has the idea of emphasizing the sense of a DM as in (11).

(11) A: il travaille pour une société de nettoyage B: **d'accord d'accord** [CODIM-ESLO] (*A: He works for a cleaning company B: **ok ok***)

In the case of semantic features, the authors also explain that in the presence of the junction "and" (e.g., **no and no**), this syndetic reduplication implies an intensive value that is different from–and even more oppositional than–simple

reduplication (e.g., **no no**). And finally, when this DM is combined with "but" (in French: *mais non !*), there are some distributional constraints because it is not possible to have a reduplication after *mais*, as shown in (12), but it is possible to have it before– *non non mais*, as in example (13).

(12) a. A: Tu as un super niveau – B: **Mais non** on parle comme ça nous [CODIM-MPF] (*A: You have a great level – B: **But no** we speak like that*)

b. ? A: Tu as un super niveau – B: **Mais non non** on parle comme ça nous [CODIM-MPF] (*A: You have a great level – B: **But no no** we speak like that*)

(13) a. A: Trop chaud trop froid ? – B: **non non mais** y a un bruit [CODIM-ESLO] (*A: Too hot, too cold ? – B: **no no** but there is a noise*)

Previous literature provides important insights concerning the range of functions of some notions of repetition. However, the notions of repetition range rather widely and no precise notions of meaning or conversational context are provided.

3 A taxonomy of Sequential Repetition

Our taxonomy is based on two basic principles. As far as form goes, we concentrate on literal repetitions (though, as we will see later, in some cases, the rules that cover certain classes apply much more widely). As far as meaning goes, our classes are motivated by 'semantic transparency'—the need to postulate semantically coherent classes.

1. **Self-repair**: in this class we include RSs which involve the speaker engaging in self-repair:

   (14) a. Anon9: It must mean that there is a er an approach that says that development in the open countryside isn't normally permissible erm **unless unless unless.** Anon9: But at least it I don't see **that that that** E two is inconsistent with the phrase development in the open countryside being strictly controlled. [BNC, J9V, L392 – L393]

2. **Confirmatory Emphasis**: this class involves cases where repetition serves to confirm a potentially surprising/controversial formulation, arising from the first element of the sequence:

(15) (Hebrew) ata mskim she hu yaxtif nun-tet larosh bemilxama she **ata omer ata omer** she kol taxlita hu huisardut (Haaretz 15/05/2025) *('You agree that he is hit with an anti-tank missile in a war that **you say you say** has the only purpose of (maintaining the government)'*

  a. Mr Hunt is the **fourth — fourth!** — Tory chancellor in the space of just four months.

  b. **Every single day. Every single day. Every single day**, we see new evidence—always humiliating, always alarming—that Donald Trump is not mentally capable of discharging the duties of the presidency. (Brad DeLong substack, May 6)

3. **Sequential**: these involve cases where repetition represents a sequence of events or actions:

(16) a. John: Of course you can't put the tails hanging down on bar ten can you. You have to put them up because the notes are on or below the middle line. **Think think think.** So you'll have to reverse the tails. [BNC, FMC, L315 – L318]

  b. Emmy: So instead of **pushing pushing pushing**, let's just see how much comes more or less naturally through advertising. ⟨ unclear⟩ [BNC, J9P, L202 – L203]

4. **Doubt Elimination**: these cases involves RSs, where the speaker intends to close discussion of an issue:

(17) a. Rod: Are you doing networking really networking you see Douglas: Well marketing marketing itself. Rod: **Yeah yeah yeah.** Douglas: Yeah yeah. Er as well as applying for er vacancies that appear in newspapers as well as er registering myself with er agencies with whom I've been in contact Rod: Okay. Douglas: er within the past. [BNC, J9Y, L711 – L715]

  b. (attested, context: discussion about who will enter bathroom first) (Greek) **piyene piyene piyene** ('You go')

  c. (attested, context: two cyclists at a junction, orthogonal to each other (French)) **Allez y, allez y, allez y** ('Go ahead')

  d. (German) Krauß[. . .] rief "Ich trinke auf den Meister." Da er offenbar Papa meint, erwidert dieser "Nein *ich* trinke auf den Meister." "**Nein, nein, nein!**" ruft Krauß "*Ich* trinke auf den Meister." *(Krauß[. . .] shouted, "I drink to the master." Since he clearly means Papa, the latter replies, "No, I drink to the master." "**No, no, no!**" shouts Krauß[. . .]. "I drink to the master.")* (*Licht Spiel*, D. Kehlmann.)

5. **Onomatopeic**: cases where RSs denote a sequence of quoted real world sounds:

(18) a. Anon1: Yeah. Anon2: Two a, a day as a single dose, strictly on an empty stomach so at least half an hour to an hour before food. Anon1: Yes okay ⟨ unclear ⟩ Anon2: for five days, **boom boom boom boom boom.** It's completely different from what you've had before. Er if it upsets your tummy let us know, but that's actually very uncommon. [BNC, GYC, L165 – L172]

6. **Shortening**: cases involving the omission of redundant or obvious speech (may also indicate that the content is trivial, unimportant, or ridiculous):

(19) a. Unknown: I assure ⟨ unclear ⟩ one area a growth of three hundred and fifty ⟨ unclear ⟩, the vast majority of whom will be actually at the production line for home care assistance **etcetera etcetera etcetera.** Surely that's going to improve the quality of, of provisions? [BNC, J43, L66 – L67]

(20) a. S: They need help getting up, going to bed, going to the loo, bathing, **blah blah blah**. J: And when you have

that context where these people need help more and more and yet you're not going to be able to supply it, isn't that almost criminal? S: I mean that's a loaded word. [BNC, KRL, L2559 – L2565]

7. **Emotive interjection**: cases where the repeated sequence involves interjection of particles expressing pleasure or displeasure

    (21) a. Unknown: Is everybody happy with that layout for that? Anon5: Yes. Anon2: Yeah! Anon4: Ecstatic! Anon3: **Mm mm mm.** Don: Mm. Anon3: Nearly. Unknown: Go on Andrew. Anon5: <laughing>:[God]! [BNC, F7F, L1301 – L1309]

    b. (Constructed) A: **aw aw aw aw**, that hurt.

4 Corpus Study

At this stage of the study, we conducted a corpus analysis using English data. We extracted examples of repetitions from the British National Corpus (BNC) (Burnard, 2000). As a result, we found $3,141$ cases of three or more word repetitions[3], and annotated 511 randomly selected cases using our proposed taxonomy. The annotation was carried out by the first two authors, both fluent in English and with a background in linguistics. We calculated Cohen's kappa to assess inter-annotator agreement and obtained a substantial agreement score of 0.704. Out of the 511 annotated cases, there were 106 instances of disagreement between the two annotators. These disagreements were subsequently resolved through discussion, during which 3 cases were identified as "Ambiguous." These ambiguous cases were excluded from the final dataset. As a result, we finalized 508 annotated examples, with the corresponding statistical results presented in Table 1.

During disagreement analysis, we identified several frequent disagreement pairs: *Doubt Elimination* versus *Confirmatory Emphasis* (19 instances), *Emotive Interjection* versus *Onomatopeic* (18 instances), *Confirmatory Emphasis* versus *Sequential* (10 instances), and *Self-repair* versus *Confirmatory Emphasis* (9 instances). These patterns suggest

---

[3]We searched for sequences repeated three times, but the results also included repetitions occurring four, five, or more times.

the need for more refined annotation guidelines and clearer category definitions to reduce ambiguity and improve consistency before we extend the annotation process to other language data.

Category	Freq.	Percentage
Emotive Interjection	237	46.65%
Doubt Elimination	137	26.97%
Onomatopeic	37	7.28%
Confirmatory Emphasis	31	6.10%
Sequential	29	5.71%
Self-repair	18	3.54%
Shortening	15	2.95%
Other	4	0.79%
Total	**508**	**100.0%**

Table 1: Frequency and Percentage of Repetition Categories in the BNC.

As shown in Table 1, our proposed taxonomy covers more than 99.0% of the total examples. We used the "Other" class to capture instances that do not fit into any of the defined categories in our taxonomy. Example 22 illustrates such a case, where the repetition involves a number. It is unclear whether this reflects the repetition of a single number or simply three instances of the digit "nine" in a telephone number.

(22) a. Nigel bell: There's a free phone number O eight hundred six two six **nine nine nine.** [BNC, HMA – L289]

Annotation results reveal that the most frequent category is *Emotive Interjection*, which accounts for 46.65% of the total sample. The second and third most frequent are *Doubt Elimination* and *Onomatopeic*, representing 26.97% and 7.28% of the sample, respectively. Other categories, such as *Confirmatory Emphasis*, *Sequential*, and *Self-repair* account for 6.10%, 5.71%, and 3.54% of the total annotations, respectively. The least frequent category is the *Shortening* class, representing 2.95% of the total annotated sample.

Although the present study focuses on annotated English data, we have also prepared a French dataset for future annotation and analysis. In the context of the CODIM project, a corpus was created using various French corpora representing different discourse genres. The oral component of the corpus contains approximately 6 million tokens. Within this subcorpus, we identified $12,667$ cases of three-word repetitions, $3,239$ cases of four-word repetitions, and 877 cases of five-word repetitions.

In the case of these corpora, the most common

repetitions are those of DM of agreement and disagreement, as shown in Table 2 :

Word	3-word	4-word	5-word
Oui	2863	773	280
Non	1139	341	140
Ouais	661	201	82
Hm	1043	774	136

Table 2: Frequency of *n-gram* repetition of French Agreement and Disagreement Discourse Markers.

5 A formal description of Repetition Sequences

In this section, we offer a grounding of our taxonomy within a dialogical view of grammar (Ginzburg and Poesio, 2016; Kempson et al., 2016; Wiltschko, 2021). We also try to characterize those words that are prone to be used in certain RS constructions. A grammar that can deal with RSs needs to be incremental, able to deal with metacommunicative meaning, emotion, and the dynamics of topic change. We assume a view of cognitive states of participants as in the framework KoS (Ginzburg, 1994; Larsson, 2002; Purver, 2006; Ginzburg, 2012). A Total Cognitive State (TCS) is formally represented in (23a). Our discussion here will be at the level of Dialogue GameBoards (DGBs), whose structure is given in (23b). Here *facts* represents the shared assumptions of the interlocutors—identified with a set of propositions; dialogue moves that are in the process of being grounded or under clarification are the elements of the *pending* list; already grounded moves are moved to the *moves* list. Within *moves*, the first element has a special status given its use to capture adjacency pair coherence, and it is referred to as *LatestMove*.[4] The current question under discussion is tracked in the *qud* field, whose data type is a partially ordered set (*poset*). *Vis-sit* represents the visual situation of an agent, including his or

her visual focus of attention (*foa*), which can be an object (*Ind*), or a situation or event. *Mood* tracks a participant's public displays of emotion, crucial for emotive interjections, and non-verbal signals such as *inter alia* laughter, smiling, and head shaking/nodding.

(23) a. $\text{TCS} =_{def} \begin{bmatrix} \text{public} & : & DGBType \\ \text{private} & : & Private \end{bmatrix}$

b. $\begin{bmatrix} \text{spkr} & : & Ind \\ \text{addr} & : & Ind \\ \text{utt-time} & : & Time \\ \text{c-utt} & : & \text{addressing(spkr,addr, utt-time)} \\ \text{facts} & : & Set(Proposition) \\ \text{vis-sit} & = & \begin{bmatrix} \text{foa} & : & Ind \vee Rec \end{bmatrix} \\ & : & RecType \\ \text{pending} & : & List(LocProp) \\ \text{moves} & : & List(IllocProp) \\ \text{qud} & : & poset(Question) \\ \text{mood} & : & Appraisal \end{bmatrix}$

Purver (2004) and Ginzburg (2012) show how to account for the main classes of clarification requests using rule schemas of the form "if u is the interrogative utterance and $u0$ is a constituent of u, allow responses that are *co-propositional*[5] with the clarification question $\text{CQ}^i(u0)$ into QUD.", where '$\text{CQ}^i(u0)$' is one of the three types of clarification question (repetition, confirmation, intended content) specified with respect to $u0$. Formally:[6]

(24)
$$\begin{bmatrix} \text{pre} & : & \begin{bmatrix} \text{MaxPENDING} = \begin{bmatrix} \text{sit} = \text{u} \\ \text{sit-type} = \text{T}_u \end{bmatrix} : \text{LocProp} \\ A = \text{u.dgb-params.spkr} : \text{IND} \\ u0 : \text{sign} \\ c1 : \text{Member}(u0, \text{u.constits}) \end{bmatrix} \\ \text{effects} & : & \begin{bmatrix} \text{MaxQUD} = \text{CQ}^i(u0) : \text{Question} \\ \text{LatestMove} : \text{LocProp} \\ c1: \text{CoPropositional(LatestMove.cont,MaxQUD)} \end{bmatrix} \end{bmatrix}$$

Since they play a role in subsequent specification, we specify two cases of what $\text{CQ}^i(u0)$ amounts to, with exemplification:

(25) a. Confirmation: $\text{CQ}^i(u0) = \lambda x$ u.cont(u1.dgb-param $\rightsquigarrow$ x) (Parameter focussing)

---

[4]In line with TTR's general conception of (linguistic) classification as type assignment—record types regiment records—propositions are construed as typing relations between records (situations) and record types (situation types), or Austinian propositions (Austin, 1961; Barwise and Etchemendy, 1987); more formally, propositions are records of type $\begin{bmatrix} \text{sit} & : & Rec \\ \text{sit-type} & : & RecType \end{bmatrix}$, true iff *sit:sit-type*. The ontology of dialogue (Ginzburg, 2012) knows two special sorts of Austinian proposition: grammar types classifying phonetic events (*Loc(utionary)Prop(ositions)*) and speech acts classifying utterances (*Illoc(utionary)Prop(ositions)*).

[5]Here *CoPropositionality* for two questions means that, modulo their domain, the questions involve similar answers: for instance 'Whether Bo left', 'Who left', and 'Which student left' (assuming Bo is a student.) are all co-propositional.

[6]Given that the signs we employ (lexical entries/phrasal rules) are construed as *types for interaction*, they refer directly to values drawn from the DGB via the field *dgb-params*.

b. A: Did Bo leave B: Bo? $\rightsquigarrow$ $?Ask, A?Leave(b)$ ("Are you asking if BO left?")

c. Intended content: $CQ^i(u0)$ = $\lambda x \text{Mean}(A,u0,x)$ (Parameter identification)

d. A: Did Bo leave, clarifying the sub-utterance 'Bo' $\rightsquigarrow \lambda x Mean(A, 'Bo', x)$ ("Who are you referring to as 'Bo'?")

In order to allow for RSs, the account sketched above for clarification questions requires one fundamental refinement, detailed in (Ginzburg et al., 2014), namely that the possibility of grounding/clarification be allowed not only at each turn boundary, but at a latency which is minimally word-by-word.

5.1 Emphasis and Self-repair

We start by considering the class we have dubbed *Confirmatory emphasis*, examples of which are repeated here as (26):

(26) a. Mr Hunt is the **fourth — fourth!** — Tory chancellor in the space of just four months.

b. **Every single day. Every single day. Every single day,** we see new evidence—always humiliating, always alarming—that Donald Trump is not mentally capable of discharging the duties of the presidency. (Brad DeLong substack, May 6)

As a consequence of uttering token w_1 of a given word/phrase, perhaps perceiving unclarity in their interlocutor, A updates QUD with the confirmation question (25a). Repetition serves as confirmation that indeed A meant what they said in w_1, and this process can of course recurse.

Similarly, we obtain a direct account of cases such as (14). Here, instead of uncertainty in the interlocutor, there is uncertainty within the current speaker, which introduces into QUD the issue (25bc). This gets instantiated as 'what did I mean with w_1', which a repetition serves as an answer to and this process can of course also recurse. Support for such an account is discussed in (Ginzburg et al., 2014; Tian et al., 2017), who discuss the pervasive occurrence of spelled out self-addressed questions of this kind in the BNC, in the London-Lund corpus (Svartvik and Quirk,

1980), and in Switchboard (Godfrey et al., 1992).[7] More precisely, they show that in self-repair self-addressed questions appropriate to the class of the sought word occur productively:

(27) a. (*anticipating a locative NP*:) No, we went out on Sat, er Sunday to erm (pause) where did we go?

b. (*anticipating an NP complement*:) He can't get any money (pause) so so he can't get erm (pause) what do you call it?

c. (*anticipating a person-denoting NP*:) But you see somebody I think it was erm what's his name?

d. (*anticipating a predicative phrase*: she's erm (pause) what is she, Indian or something?
(Examples (73) in (Ginzburg et al., 2014))

These are sub-questions of the issue 'what is the word I mean to utter', hence licensed by (25c).

These class of cases, then, can be explicated using means postulated in previous work on self/other-repair.

5.2 Doubt Elimination

This class we suggest involves a basic contextual effect: it involves a question under discussion as a precondition, and the effect is the indication by the speaker that the question is no longer at issue. For instance, for the cases in (28) we have QUDs respectively 'who will enter the bathroom first' and 'who should move first from the junction:

(28) a. (attested) (Greek) **piyene piyene piyene** ('You go')

b. (attested, context: two cyclists at a junction, orthogonal (French)) **Allez y, allez y, allez y** ('Go ahead')

We break this in two: a conversational move of doubt elimination, given in (29a) and a construction involving n copies of an utterance, given in (29b). The construction has as its n daughters, n phonologically identical proposition–denoting elements, and yields as content an utterance whose force is *doubt elimination*:[8]

---

[7]We thank an anonymous reviewer for raising this issue.

[8]The formulation of (29b) involves certain formal challenges, since it constitutes a schema over n identical daughters, with $n \geq 2$, but we will leave making this precise to another occasion.

(29) a. **DoubtElimination:**

$$\left[\begin{array}{l} \text{pre:} \begin{bmatrix} \text{QUD} = \langle q,Q \rangle : \text{poset(Question)} \\ \text{LatestMove} = \text{DoubtElim(p,q)} \end{bmatrix} \\ \text{effect:} \begin{bmatrix} \text{QUD} = Q : \text{poset(Question)} \end{bmatrix} \end{array} \right]$$

b. *doubt-elim-cl* =

$$\left[\begin{array}{l} \text{dgb-params:} \begin{bmatrix} q : \text{Question} \\ \text{QUD} = \langle q,Q \rangle : \text{poset(Question)} \\ u0 : \text{Rec} \\ \text{spkr: IND} \\ \text{addr: IND} \\ \text{utt-time: TIME} \\ c1 : \text{addressing(spkr,addr,utt-time)} \\ c2: \text{resolve(p,q)} \\ c3: \text{member(p, FACTS)} \end{bmatrix} \\ \text{cont} = \begin{bmatrix} \text{sit} = u0 \\ \text{SIT-TYPE} = \begin{bmatrix} R = \text{DoubtElim} : \text{IllocReln} \\ c1 : R(\text{spkr,addr,utt-time,p,q}) \end{bmatrix} \end{bmatrix} : \text{Prop} \\ \qquad\qquad | \\ \begin{bmatrix} \text{dtrs} = \langle X_1, X_2, \ldots, X_n \rangle : \text{list(sign)} \ (n \geq 2) \\ c1 : =(X_1.\text{phon}, X_2.\text{phon}, \ldots, X_n.\text{phon}) \\ p = X_1.\text{cont} : \text{Prop} \end{bmatrix} \end{array} \right]$$

We can use the rules in (29) to explain some cases of infelicity of RS, brought up in section 1: RS (as *doubt elimination* use) is infelicitous in a courtroom Q/A since it is not up to a witness to eliminate a question from discussion. RS is infelicitous in a backchannel of an incomplete utterance since the backchanneler does not have the right to eliminate the issue being introduced by the speaker.

The rule in (29b) might seem idiosyncratic. However, it resembles *formally* rules that need to be postulated for semantically related head gestures like shakes and nods (Wagner et al., 2014). For instance, using the notational system of (Crasborn, 2014), Lücking and Ginzburg (2023) postulate (30a) as a lexical entry for a head shake, whose content is identical to 'No', whereas (30b) as a noetic head shake (a shake whose content involves valenced amazement). The former involves an underspecified number of shakes, the latter $n \leq 3$ at a slow rate. An analogous specification would be needed to distinguish different types of nods (Hadar et al., 1985).

(30) a.
$$\begin{bmatrix} \text{phon : no/shape : Shake-n} \\ \text{content = Assert(spkr,addr,} \\ \text{u-time,NoSem}(p)) : \textit{IllocProp} \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{shape : Shake-slow-3} \\ \text{cont = Pos-amaze(spkr,p,}\delta) : \textit{Prop} \end{bmatrix}$$

5.3 Utterance redundancy signalling

A lexical entry for a word like 'etc' is given in (31): this is, of course, an intrinsically incremen-

tal meaning. It indicates that the *projected* content (Ginzburg et al., 2020; Cooper, 2023) of the utterance is unnecessary:

(31)
$$\begin{bmatrix} \text{phon : etsetra} \\ \text{dgb-params :} \begin{bmatrix} \text{proj.sit-type.cont =} \\ a : \text{SemObj} \end{bmatrix} \\ \text{cont = UnNecessary}(a) : \text{Prop} \end{bmatrix}$$

5.4 Pleasure expression

We assume a pleasure interjection like 'mm' has a lexical entry as in (32). It does not have propositional content, in contrast to e.g., laughter (Mazzocconi et al., 2020) and agreeing with (Kaplan, 1999). This force serves as the trigger for the conversational rule in (33) (Ginzburg and Kim, 2023), which updates the DGB Mood, depending on the valence of the exclamation. Thus, a sequence of such exclamations gives rise to sequential signalling of increasing/decreasing pleasantness, depending on the valence of the exclamation.

(32)
$$\begin{bmatrix} \text{dgb-params :} \begin{bmatrix} \text{spkr : Ind} \\ \delta : \text{Degree} \end{bmatrix} \\ \text{cont = EmotInt(spkr,}\delta) \end{bmatrix}$$

(33) ScaleUp-based exclamation

a.
$$\left[\begin{array}{l} \text{tcs=} \begin{bmatrix} \text{dgb : DGBType} \\ \text{private : Private} \end{bmatrix} : \text{TCS} \\ A = \text{dgb.spkr : IND} \\ A. \text{ preconds:} \begin{bmatrix} v = \text{valence(EmotInt)} : \text{Boolean} \\ \delta : \text{Degree} \\ \text{LatestMove.cont =} \\ \text{EmotInt(spkr,}\delta) : \\ \text{IllocProp} \end{bmatrix} \\ A.\text{effect :} \begin{bmatrix} \text{PolPleasantnessIncr}(\delta, \epsilon) \end{bmatrix} \end{array} \right]$$

b. PolPleasantnessIncr(v) =
$$\left\{ \begin{array}{ll} PositivePleasantnessIncr & if \ v : + \\ NegativePleasantnessIncr & if \ v : - \end{array} \right\}$$

c. PositivePleasantnessIncr$(\delta, \epsilon) =_{def}$
$$\left[\begin{array}{l} \text{preconditions:} \begin{bmatrix} \text{LatestMove.cont : IllocProp} \end{bmatrix} \\ \text{effect :} \begin{bmatrix} \text{Mood.pleasant.arousal.pve =} \\ \epsilon(\text{preconds.Mood.pleasant.arousal.pve}) + (1 - \epsilon)\delta : \text{Real} \\ \text{Mood.pleasant.arousal.nve =} \\ \epsilon(\text{preconds.Mood.pleasant.arousal.nve}) : \text{Real} \end{bmatrix} \end{array} \right]$$

d. NegativePleasantnessIncr$(\delta, \epsilon) =_{def}$
$$\left[\begin{array}{l} \text{preconditions:} \begin{bmatrix} \text{LatestMove.cont : IllocProp} \end{bmatrix} \\ \text{effect :} \begin{bmatrix} \text{Mood.pleasant.arousal.nve =} \\ \epsilon(\text{preconds.Mood.pleasant.arousal.nve}) + (1 - \epsilon)\delta : \text{Real} \\ \text{Mood.pleasant.arousal.pve =} \\ \epsilon(\text{preconds.Mood.pleasant.arousal.pve}) : \text{Real} \end{bmatrix} \end{array} \right]$$

6 Discussion: the idempotence constraint

The final issue we consider is which words are susceptible to RS. Of course, to a first approximation, all words are equally prone to give rise to self-repair and to confirmatory emphasis (though of course the former is driven by ease of lexical access, the latter perhaps by surprisal potential, which is heavily context dependent.). The onomatopeic and sequential classes are fairly randomly conditioned by newsworthy sounds occurring in the world and the existence of repetitive events, respectively. Emotive interjections are caused by non-instantaneous events, so will remain as triggers for a certain period. This leaves **Doubt Elimination** and **shortening**. There is a tempting explanation for such uses in that the meanings posited for such words are what one might call *idempotent*—in a sequence exemplified in (34a) the content assigned to w_1 creates a context appropriate for w_i ($i > 1$), which leads to content identical to the content of w_1 (putting aside the effects of the rule **Doubt Elimination**). 'Yes' requires a positive polar question $p?$ as MaxQUD and outputs p as its content, whereas 'no' outputs a negative proposition as its content, and when it has a negative polar question as input $\neg p?$ outputs $\neg p$ as content.

(34) a. $w_1 w_2 \ldots w_n$

     b. Meaning of 'yes':
$$\begin{bmatrix} \text{MaxQUD} = p? : \text{PosPolarQuestion} \\ \text{cont} = p : \text{PosProp} \end{bmatrix}$$

     c. Meaning of 'no':
$$\begin{bmatrix} \text{MaxQUD} = p? : \text{PolarQuestion} \\ \text{cont} = \text{NoSem}(p) : \text{NegProp} \end{bmatrix}$$

     d. Content of 'yes yes ... yes': $pp \ldots p$

     e. Content of 'no no no': $\neg p \neg p \ldots \neg p$

The French word 'si' also occurs productively as an RS:

(35) a. A: C'est vrai qu'il faut dire les choses. B: Mais je pense que les choses elles ont été claires depuis le début. A: Si si si si si si si.

     b. A: vous faites euh des activités de loisir vous n'en faites pas A: si ? B: si si si

Is 'si' idempotent? On the formulation in (36a), it is not: it requires a negative proposition/polar question as its immediate left context and outputs a positive proposition (Noveck et al., 2021; Abeillé and Godard, 2021). However, another view of 'si' is possible, as given in (36b)—it involves a double negative. On this view, 'si' is idempotent.

(36) a. $$\begin{bmatrix} \text{MaxQUD} = \neg p? : \text{NegPolarQuestion} \\ \text{cont} = p : \text{PosProp} \end{bmatrix}$$

     b. $$\begin{bmatrix} \text{MaxQUD} = \neg p? : \text{NegPolarQuestion} \\ \text{cont} = \neg\neg p : \text{NegProp} \end{bmatrix}$$

Finally, we note that words like 'etc' and 'blah' are also idempotent: they indicate that their right context can be omitted, is predictable, etc. This recurses to the right.

7 Conclusions and Future Work

Disallowing non-modifier words and phrases to occur in duplicate copies is a basic principle of formal grammar. The existence of such sequences is pervasive in conversation, with a wide variety of meanings, across diverse languages: our main data sources in this paper are English and French, but we provide also examples from German, Hebrew, and Chinese—we hypothesize that the phenomenon is universal and related to its even more pervasive occurrence in head gestures, across all three axes (tilting, nodding, shaking). Our cross-linguistic hypothesis needs testing.

Although the existence of the phenomenon has been noted in previous literature, no systematic taxonomy nor formal treatment has been offered. In this paper, we offer both a taxonomy with high coverage and sketch the meanings RSs can bear in a dialogical grammar. These meanings, of course, bear refinement, as do the different prosodic realizations of the different classes. We also offer a tentative characterization of the words that are particularly susceptible to several classes of RS uses.

References

Anne Abeillé and Danièle Godard. 2021. *La grande grammaire du français*. Éditions Actes Sud.

John L. Austin. 1961. Truth. In James Urmson and Geoffrey J. Warnock, editors, *Philosophical Papers*. Oxford University Press. Paper originally published in 1950.

Jon Barwise and John Etchemendy. 1987. *The Liar*. Oxford University Press, New York.

Wolfram Bublitz. 1989. Repetition in spoken discourse. *Anglistentag*.

L. Burnard. 2000. *Reference Guide for the British National Corpus (World Edition)*. Oxford Universtity Computing Services.

Robin Cooper. 2023. *From Perception to Communication: a Theory of Types for Action and Meaning*. Oxford University Press.

Onno Crasborn. 2014. Annotating form + function of head movements. http://signnonmanuals.aau.at/sites/default/files/pdf/Crasborn_SignNonmanuals_Klagenfurt_2014_headmovement.pdf.

Gaétane Dostie. 2007. La réduplication pragmatique des marqueurs discursifs. de là à là là. *Langue française*, 154(2):45–60.

Gaétane Dostie. 2011. La réduplication en langue et en discours. oui oui oui, non non non non non. In Sylvie Hancil, editor, *Le rôle de l'affect dans les marqueurs discursifs*, pages 133–154. Presses universitaires de Rouen et du Havre, Mont-Saint-Aignan.

Frederick Erickson. 1984. Rhetoric, anecdote, and rhapsody: Coherence strategies in a conversation among black american adolescents. In Deborah Tannen, editor, *Coherence in Spoken and Written Discourse*, pages 91–102. Ablex, Norwood, NJ.

Franck Floricic and Françoise Mignon. 2007. Non non et no no en français et en italien : réduplication ou réitération ? *Faits de Langues*, (29):49–61.

Jonathan Ginzburg. 1994. An update semantics for dialogue. In H. Bunt, editor, *Proceedings of the 1st International Workshop on Computational Semantics*. ITK, Tilburg University, Tilburg.

Jonathan Ginzburg. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford University Press, Oxford.

Jonathan Ginzburg, Robin Cooper, Julian Hough, and David Schlangen. 2020. Incrementality and hpsg: Why not. In *Constraint-based syntax and semantics: Papers in honor of Danièle Godard. CSLI Publications*. CSLI Publications.

Jonathan Ginzburg, Raquel Fernández, and David Schlangen. 2014. Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics*, 7(9):1–64.

Jonathan Ginzburg and Jong-Bok Kim. 2023. Exclaiming non-sententially: exclamative sluices and the (non-elliptical) nature of ellipsis constructions. *Glossa: a journal of general linguistics*, 8(1).

Jonathan Ginzburg and Massimo Poesio. 2016. Grammar is a system that characterizes talk in interaction. *Frontiers in Psychology*, 7:1938.

John J. Godfrey, E. C. Holliman, and J. McDaniel. 1992. Switchboard: Telephone speech corpus for research and devloment. In *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, pages 517–520, San Francisco, USA.

Uri Hadar, Timothy J Steiner, and F Clifford Rose. 1985. Head movement during listening turns in conversation. *Journal of Nonverbal Behavior*, 9(4):214–228.

Fuhui Hsieh. 2011. Repetition in social interaction: A case study on mandarin conversations. *International Journal on Asian Language Processing*, 19(4):153–168.

Hans Kamp, Josef Van Genabith, and Uwe Reyle. 2010. Discourse representation theory. In *Handbook of Philosophical Logic: Volume 15*, pages 125–394. Springer.

David Kaplan. 1999. The meaning of ouch and oops: Explorations in the theory of meaning as use. *Manuscript, UCLA*.

Ruth Kempson, Ronnie Cann, Eleni Gregoromichelaki, and Stergios Chatzikyriakidis. 2016. Language as mechanisms for interaction. *Theoretical Linguistics*, 42(3-4):203–276.

Keith T Kernan. 1977. Semantic and expressive elaboration in children's narratives. In *Child discourse*, pages 91–102. Elsevier.

Staffan Larsson. 2002. *Issue based Dialogue Management*. Ph.D. thesis, Gothenburg University.

Andy Lücking and Jonathan Ginzburg. 2023. Leading voices: Dialogue semantics, cognitive science, and the polyphonic structure of multimodal interaction. *Language and Cognition*, 15(1):148–172.

Chiara Mazzocconi, Ye Tian, and Jonathan Ginzburg. 2020. What's your laughter doing there? a taxonomy of the pragmatic functions of laughter. *Transactions of Affective Computing*.

Neal R Norrick. 1987. Functions of repetition in conversation. *Text-Interdisciplinary Journal for the Study of Discourse*, 7(3):245–264.

Ira Noveck, Nicolas Petit, Ye Tian, and Giuseppina Turco. 2021. Revealing pragmatic processes through a one-word answer: When the french reply si. *Journal of Memory and Language*, 120:104245.

Matthew Purver. 2004. *The Theory and Use of Clarification in Dialogue*. Ph.D. thesis, King's College, London.

Matthew Purver. 2006. Clarie: Handling clarification requests in a dialogue system. *Research on Language & Computation*, 4(2):259–288.

Ghaleb Rabab'ah and Ali Farhan AbuSeileek. 2012. The pragmatic functions of repetition in tv discourse. *Research in Language*, 10(4):445–460.

Tanya Stivers. 2004. "no no no" and other types of multiple sayings in social interaction. *Human Communication Research*, 30(2):260–293.

J. Svartvik and R. Quirk. 1980. *A Corpus of English Conversation*. CWK Gleerup.

Deborah Tannen. 1989. *Talking voices: Repetition, dialogue, and imagery in conversational discourse*, volume 26. Cambridge University Press.

Ye Tian, Takehiko Maruyama, and Jonathan Ginzburg. 2017. Self addressed questions and filled pauses: A cross-linguistic investigation. *Journal of psycholinguistic research*, 46(4):905–922.

Petra Wagner, Zofia Malisz, and Stefan Kopp. 2014. Gesture and speech in interaction: An overview.

Martina Wiltschko. 2021. *The grammar of interactional language*. Cambridge University Press.