

1.1 The absence of the grammaticality effect

What is arguably of greater *theoretical* interest are those cases where the sentence-matching task apparently ‘fails’: that is, where superficially ungrammatical constructions are treated in terms of response latency as though they were grammatical. These results have been taken by some researchers as evidence of the ‘underlying grammaticality’ of such constructions, especially where there is independent theoretical evidence to support this claim; see, for example, Freedman & Forster (1985), Forster & Stevenson (1987), Clahsen & Hong (1995). This is also the case in the present experiment.¹

One such construction, studied in Duffield & White (1999) and Duffield et al (2002), is what may be termed the ‘*medial French causative’ (*MFC), where an accusative (subject) clitic appears in medial position between the higher causative verb and the lower predicate with which it is thematically associated, as in (2a). With respect to the surface grammar, this structure is ungrammatical; however, in previous SMT experiments, *MFCs were treated *by native-speakers* no differently from their grammatical counterpart (2b), and in clear contrast to comparable ungrammatical pairs, as exemplified in (2c), which elicited reliably longer response latencies.

- (2) a. *Je fais *le* voir.
I make him see-inf.
‘I make him see.’
- b. Je *le* fais voir.
I him make see-inf.
‘I make him see.’
- c. *Je fais voir *le* aujourd’hui.
I make see him today
‘I make him see (today).’

¹ A reviewer asked whether there are any specific properties (i.e., categories, features, or operations) common to all constructions that failed to show grammaticality effects in SMTs. In pre-Minimalist vocabulary, the answer is affirmative (if somewhat loose): sentences may be underlyingly grammatical if they involve a non-surface grammatical level of representation—e.g., LF (post reconstruction), or s-structure where Operator-Binding takes place at LF—that differs from the surface representation. It is harder to capture this in a purely derivational framework.

The absence of a grammaticality effect in this case could be dismissed as a task or item effect, were it not for the following considerations. First, in all other comparable conditions in the same experiment, native-speakers’ results yielded strong, reliable effects of grammaticality. This was true, for example, of contrasting pairs of sentences involving restructuring verbs such as *pouvoir*, *vouloir*, etc., illustrated in (3).²

- (3) a. Je veux *le* voir.
I want him see-inf.
‘I want to see him.’
b. *Je *le* veux voir.

Second, the same subjects’ judgments of the relevant items in a subsequent (traditional) offline GJ task showed a uniform rejection of these *MFC sentences, showing that the absence of a grammaticality effect in the SM task was not due to poorly chosen stimuli.

Third, experiments using separate materials and a different instrument—a self-paced reading task—have shown independently that this construction is rather marked in processing terms. Dwivedi & Hoover (1996), and Hoover & Dwivedi (1998), have shown that readers prefer to associate a clitic with the verb that immediately follows it, a preference which yields longer processing times³ for (grammatical) initial French causatives (2b) over comparable causative sentences without clitics. We return to this point directly.

Fourth—the focus of the earlier papers—L2 learners did in fact show reliable grammaticality effects for the *MFC condition; in this respect, they appeared to outperform native-speakers.

Finally, as mentioned earlier, there are theoretical grounds for considering the *MFC construction to be underlyingly grammatical. Most recent generative analyses of Romance causative constructions assume that French *faire* + *infinitive* constructions remain fundamentally biclausal, where the lower clause contains all

² Note that here the expected judgments are reversed: medial placement of the clitic is grammatical, initial placement is ungrammatical.

³ At the point of presentation of the second verb.

of the functional architecture of a normal matrix clause, including clitic projections, and where the surface (initial) position of the clitic is derived by A'-movement from this lower (medial) clitic projection. This derivation is illustrated in (4) below. For relevant discussion see Sportiche (1996), Roberts (1997), Guasti (1993), Guasti (1996); also Vinka (2002).

(4) [je le_i fais [... [CIP-Acc le_i ... [VP pro_i voir]]]]

1.1.1 Gass' (2001) criticisms

Notwithstanding the preceding discussion, the SMT has been vulnerable to criticism on various counts. Early critics, notably Crain & Fodor (1987), called for a radical reinterpretation of the original findings with native-speakers, claiming that a low-level factor of 'correctability' better explained the early data than any appeal to more interesting theoretical notions such as underlying grammaticality.⁴

More recently, Gass (2001) has criticized many of the previous experiments involving L2 learners on separate grounds, both methodological and empirical. Gass' principal criticisms are based on the observed contrast between previous work and her own experiments using SM. Those contrasts may be briefly summarized. First, in Gass' own initial experiment, L2 learners showed no reliable grammaticality effect in *any* structural condition (adverb placement, subject-verb agreement, clitic placement), nor were grammatical sentences always matched faster than ungrammatical sentences by native-speakers. Second, in a follow-up experiment, intermediate learners of French did not perform significantly differently from beginning learners who—according to Gass—could not be presumed able to construct any higher-level structural representations whatsoever. Third, the L2 results from Gass' first experiment did not reliably correlate with the same subjects' offline judgments of grammaticality (indeed, a small but reliable 'reverse grammaticality effect' was observed). Finally, Gass

⁴ Crain & Fodor's position on SM is uncompromisingly harsh: '...this overgeneration account of the matching task results is incorrect [:] the relevant difference between the two sentence types is whether the ungrammaticality tends to be spontaneously corrected by subjects. This correctability account of the data is of *absolutely no theoretical interest*, but unfortunately it appears to be correct (Crain & Fodor 1987:123-124) [our italics].' An equally firm reply to Crain & Fodor's account is provided by Forster & Stevenson (1987).

criticized certain aspects of the experimental design of previous experiments, including choice of materials and timing issues.

Setting aside for a moment the non-trivial methodological issues, Gass’ other criticisms reduce to two essential points: first, in Gass’ own studies, the SMT failed to yield any reliable grammaticality effects, casting doubt on the replicability of earlier results; second, mismatches were observed between subjects’ results in offline GJ tasks vs. those on the SMT, again suggesting (to Gass) that SM does not provide a reliable measure of grammaticality.

With respect to the first point, it is quite clear that experimental results are only valid to the extent that they are replicable: the main purpose of the present set of experiments is to show such replicability. However, it is equally clear that failure to replicate on any given occasion does not by itself invalidate an earlier finding. We will not speculate here on the reasons for the failure of SM in Gass’ experiments. What is crucial to notice, however, is that her experiments L2 learners’ results failed *globally* to show any reliable grammaticality effect; by contrast, in Duffield *et al.* (2002), and several other papers, the SM task elicited reliable grammaticality effects for L2 learners and native speakers overall and in most subconditions: it was only in *one specific subcondition* (*MFC) that native speakers’ results failed to yield a reliable effect.⁵

The validity of the second point, concerning the mismatches between online and offline results in Gass’ experiments, rests on the premise that offline acceptability judgments provide a reliable litmus test of grammaticality. Without wishing to belabour the point here—see Duffield (2003, 2004) for extended discussion—we reject this premise. In fact, if it were valid, there would be no real point in using other instruments such as SM to assess grammaticality, since whenever a systematic divergence were found between results, the traditional GJ results would immediately be deferred to (as is the case in Gass’ paper).

⁵ Such cases of local failure are much more profitably viewed in terms of a (statistically) *reliable interaction* between grammaticality and condition than in terms of an *unreliable main effect* of grammaticality (as Gass would have it).

If this brief response is adequate, what remains from Gass' paper is an even greater urgency to successfully replicate the earlier results, and to develop an experimental design that allows us to better understand the grammaticality effect in SM. With this in mind, we conducted two follow-up experiments: the first, a near copy of the restructuring and causative sub-conditions of Duffield *et al*'s larger experiment; the second, an extension of the causative conditions, using TRANSITIVE causatives, contrasting *medial-accusative with *medial-dative clitics, as exemplified in (5) and (6) below:

- (5) a. Je *la* fais suivre par son frère.
I her make follow by her brother
'I make her brother follow her.'
- b. *Je fais *la* suivre par son frère.
- (6) a. On *lui* fait regarder la télévision.
one her makes watch the television
'She is made to watch television.'
- b. *On fait *lui* regarder la télévision.

Before discussing these experiments, it is necessary to address an important—but hitherto unclarified—methodological issue, namely, the locus of the (absent) effect.

1.1.2 The locus of the effect

In the SMT, the grammaticality effect (GE) is operationally defined as the difference between the mean response latency for grammatical items in a particular condition, and that for the corresponding ungrammatical items ($GE = RT_{UGR} - RT_{GR}$). In principle, the absence of such an effect in medial causatives might come about in one of two ways: either because of some processing difficulty with the *grammatical* pairs, producing inhibition, or because of relatively greater ease of processing with the *ungrammatical* ones, producing facilitation.

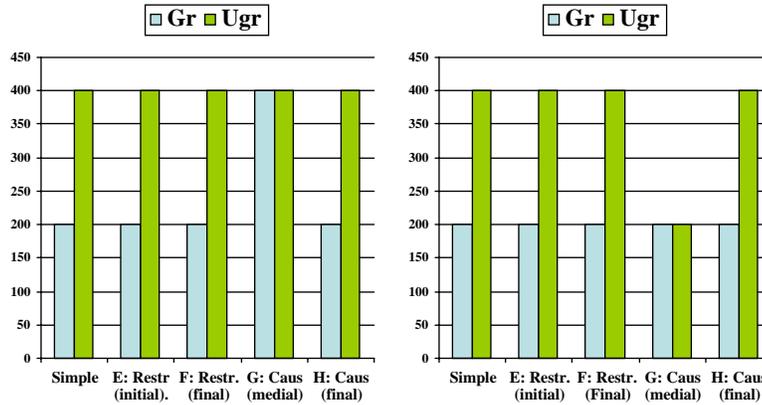


Figure 1. Contrasting Effects of Inhibition and Facilitation

In the ideal case, if the absence of any GE were due to inhibition, we would expect to see *longer response latencies* for grammatical pairs for those forms vis-à-vis other comparable *grammatical* pairs (Fig. 1: LH diagram); conversely, if this were due to facilitation of the ungrammatical pairs, we would expect to see *shorter latencies* for these vis-à-vis other comparable *ungrammatical* pairs (Fig. 1: RH diagram). Clearly, only the latter result supports the idea of underlying grammaticality. However, the results of the Hoover & Dwivedi study mentioned above would lead us to expect an inhibition-based account instead.

In Duffield *et al.* (2002), native speakers’ results showed both inhibition for the grammatical causatives (-37 msecs) *and* facilitation for the ungrammatical medial causatives (+51 msecs), *when compared to the overall means* (Fig. 2: LHS). However, when the results from medial causatives—cf. (2a)/(2b) above—were compared, more locally, with those of ‘final’ causatives —(2b/2c)—only a strong facilitation effect (+140 msecs) was found (Fig. 2: RHS). This offers some *a priori* support for the underlying grammaticality interpretation.

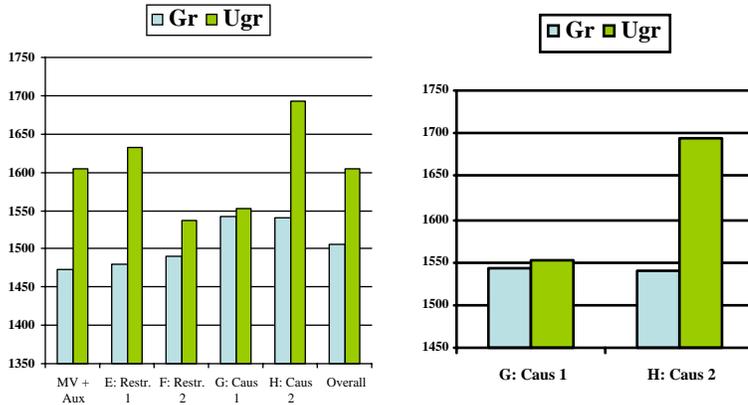


Figure 2: Broad vs. Narrow Range Comparisons

2. Experiments

2.1 Design

We conducted two experiments, using the same experimental set-up reported in Duffield *et al.* (2002) but with some modifications in the materials with respect to: the number of separate conditions in each experiment (reduced); the proportion of fillers/distracter items to test items (reduced); the number of test items per condition (increased); and the number of subjects tested (increased). We also paid particular attention to balancing the gender/number properties of the pronominal clitics in the test items {*le, la, les; lui, leur*}.

Both experiments manipulated the (surface) grammaticality of the pairs, the construction type involved, the position of the clitic with respect to the two verbs {initial, medial (final)}, and the number and gender of the clitic {masc., fem.; singular, plural}.

Experiment I re-examined the contrast between causative and restructuring verbs, previously exemplified in (2) and (3) above. Experiment II probed a contrast *within* the class of causative constructions: in contrast to Experiment I, which involved an INTRANSITIVE lower predicate, Experiment II examined clitic placement with two classes of TRANSITIVE causatives, contrasting the *faire + infinitive* construction with the so-called *faire par* construction, as was illustrated in (5) and (6) above. Though converging with respect to the grammatical (initial)

placement of the clitic, these two constructions are distinguished by the morphological case and understood grammatical relation of the clitic pronoun. In the *faire par* construction, the clitic is interpreted as the *object* of the lower clause and realized with accusative case (just like *subject* pronouns of intransitive lower verbs discussed earlier). By contrast, in the transitive *faire + infinitive* construction in (6), the clitic is interpreted as the lower clause *subject* but is realized with dative, rather than accusative case.

Investigating transitive causatives enables us to tease apart two possible interpretations of facilitation in *MFC constructions, supposing this (non)-effect is replicated. A grammatical interpretation of the *MFC effect would predict that pairs containing *medial *accusative* clitics should be treated as underlyingly grammatical, independently of their grammatical relation (subject/object). An alternative ‘shallow syntactic processing’ account—along the lines proposed in Clahsen & Felser (forthcoming)—might predict that medial clitic placement would be accepted only where the order is congruent with canonical thematic relations (SUBJ. (AGENT)₁-VERB₁-SUBJ. (AGENT)₂-VERB₂): that is to say, it predicts that (2b) and (6b) pairs should pattern together, with (5b) pairs showing a significant GE.

2.1 Subjects

We tested 35 adult native-speakers of French at two university sites in Canada and Europe.

2.1 Results

Data from 29 subjects—pending statistical analysis—reveals the following.⁶ First with respect to Experiment I, Fig. 3 (LHS) shows that the core result of previous experiments appears to have been entirely replicated. In contrast to the grammaticality effects observed in both restructuring conditions (EF), as well as in the (immediately comparable) final causative condition (H), absolutely no GE is observed in the crucial *MFC condition (G); indeed, the interaction between

⁶ Results from five subjects were excluded due to high error rates in matching, or exceptionally slow mean response latencies. One test item (from restructuring condition F) was also excluded. The final analysis was based on the response latencies of the remaining subjects/items that fell within 2 standard deviations of each subject’s individual overall mean for correctly matched items.

clitic position and construction type seems even sharper than before. Moreover, the most immediate comparison between causative conditions G and H once more suggests the absence of a GE as due to facilitation of the ungrammatical pairs in G—as predicted by the underlying grammaticality account—rather than to inhibition of the grammatical pairs.

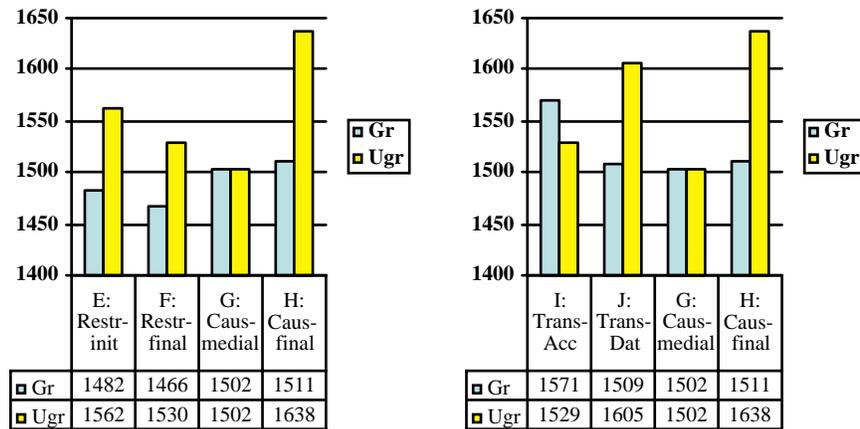


Figure 3: Preliminary Results: Experiments I & II

The results of Experiment II, with transitive causatives, are somewhat harder to interpret at this point, given that the accusative condition (I) appears to show not only (predicted) facilitation for the ungrammatical forms but also a surprising inhibition of the grammatical pairs. This obviously requires further investigation. However, the main result here is the contrast between the accusative condition (I), where no GE is observed, and the dative condition (J), which apparently shows a robust grammaticality effect. This suggests that case, rather than thematic roles or surface grammatical relations, determines the acceptability of medial causatives; if so, this speaks against any shallow processing explanation of the result, and is more consistent with the grammatical account developed here.

Contra Gass (2001), SM can tell us something interesting and valuable about competence grammar and competent parsing.

References

- Birdsong, David (1989). *Metalinguistic performance and interlinguistic competence*. New York: Springer-Verlag.
- Bley-Vroman, Robert; and Masterson, Deborah (1989). Reaction time as a supplement to grammaticality judgments in the investigation of second language competence. vol. 8.2. *University of Hawai'i Working Papers in ESL*. 207-237: University of Hawai'i.
- Clahsen, Harald; and Hong, Upyong (1995). Agreement and null subjects in German L2 development: new evidence from reaction-time experiments. *Second Language Research* 11, 57-87.
- Clahsen, Harald; and Felser, Claudia (forthcoming). Grammatical processing in language learners. *Journal of Applied Psycholinguistics*.
- Crain, Stephen; and Fodor, Janet Dean (1987). Sentence matching and overgeneration. *Cognition* 26, 123-169.
- Duffield, Nigel (2003). Measures of Competent Gradience. In *The Lexicon-Syntax Interface in Second Language Acquisition*, Roeland Van Hout, Aafke Hulk, Folkert Kuiken and Richard Towell (eds.), 97-127. Amsterdam & Philadelphia: John Benjamins Publishing Company.
- (2004). Implications of Competent Gradience. *Moderne Sprachen* 45 (1), 95-117.
- Duffield, Nigel; and White, Lydia (1999). Assessing L2 knowledge of Spanish clitic placement: converging methodologies. *Second Language Research* 15 (2), 133-160.
- Duffield, Nigel; White, Lydia; Bruhn De Garavito, Joyce; Montrul, Silvina; and Prévost, Philippe (2002). Clitic placement in L2 French: evidence from sentence matching. *Journal of Linguistics* 38 (3), 1-37.
- Dwivedi, Veena D; and Hoover, Michael L (1996). Processing the French causative construction. McGill University: ms.
- Forster, Kenneth (1979). Levels of processing and the structure of language processor. In *Sentence Processing*, W.E. Cooper and E.C.T. Walker (eds.). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Forster, Kenneth I; and Stevenson, B (1987). Sentence-Matching and well formedness. *Cognition* 26, 171-186.
- Freedman, Sandra E.; and Forster, Kenneth I. (1985). The psychological status of overgenerated sentences. *Cognition* 19, 101-131.
- Gass, Susan M (2001). Sentence Matching: a re-examination. *Second Language Research* 17 (4), 421-441.
- Guasti, Maria-Teresa (1993). *Causative and perception verbs*. Turin: Rosenberg and Sellier.
- (1996). Semantic restrictions in Romance Causatives and the Incorporation Approach. *Linguistic Inquiry* 27, 294-312.
- Hoover, Michael L; and Dwivedi, Veena D (1998). Syntactic processing by skilled bilinguals. *Language Learning* 48 (1), 1-29.
- Roberts, Ian (1997). Restructuring, head movement and locality. *Linguistic Inquiry* 28, 423-460.
- Sportiche, Dominique (1996). Clitic constructions. In *Phrase structure and the lexicon*, Johann Rooryck and Laurie Zaring (eds.), 213-276. Dordrecht: Kluwer.
- Vinka, Mikael (2002). Causativization in North Sami. Department of Linguistics, McGill University: Doctoral dissertation.