

BLEND RECOGNISABILITY IN ENGLISH AS A FOREIGN LANGUAGE An experiment

ELISA MATTIELLO
UNIVERSITÀ DI PISA

Abstract – Recognisability is one of the major constraints that most linguists place on lexical blends and their well-formedness. Blends indeed display an unpredictable output that is not transparently analysable into morphemes, and their source words are difficult to recognise for both hearers and readers. The possible combinatory patterns of the source lexemes, the different portions that are retained in the final blend, and their semantic contribution to the overall meaning increase the number of variables and classificatory criteria for blends, thus decreasing predictability of the output given an input. For students of EFL, lexical blends are even more difficult to access due to the fact that the language in which they are formed is not their native language. This paper reports on results from an experiment on 18 Italian students who were tested on English blends. The participants were asked to identify the source words and meanings of a number of blends selected according to different (phonological, morphotactic, semantic) criteria. The results of the experiment show that the recognisability of English lexical blends by Italian native speakers depends on 1) the type of characteristics that the blend displays (overlap between the source words, semantic weight of the source words, headedness, same prosodic structure as one of the source words), 2) the category (substitution vs. overlap, coordinate vs. attributive) to which the blend belongs, and 3) the context where it is used. In general, the experiment sheds some light on the type of processes (e.g., decomposition and textual reference) involved in the recognition and accessibility of English lexical blends.

Keywords: lexical blends; EFL; recognisability; source words.

1. Introduction

Lexical blends are words formed by merging parts of two or more source words (henceforth SWs), with partial loss of at least one of them, and often with phonemic or graphemic overlap. Commonly cited examples of English blends are *smog* ← *sm(oke)* + (*f*)*og* and *brunch* ← *br(eakfast)* + (*l*)*unch*. The loss or fusion of phonemic/graphemic material inevitably makes the process of blend accessibility more complicated to the language speaker and of even more difficult comprehension to the non-native speaker.

Several studies demonstrate that, in order to be well-formed and successful, both elements in the blend must be recognisable (Bat-El 2006; Bauer 2012; Cannon 1986; Gries 2004; Mattiello 2013). This is generally

known as the Principle of Maximisation, i.e., to preserve as much of the SWs as it is optimal for their recognisability (Beliaeva 2014; Gries 2004, 2006). On the other hand, the Principles of Least Effort (Zipf 1949) and of Linguistic Economy (Martinet 1955) favour the formation of new blend words that are brief, not redundant, easy to pronounce, and pleasing to the ear. Accordingly, new English blends should be created as a compromise between maximisation of segments from the SWs and minimisation of efforts in production, perception, and recognition.

Investigating the individual quantitative contributions of SW₁ and SW₂ to blends, Gries (2004, p. 664) has demonstrated the greater relevance of SW₂ compared with SW₁ in terms of number of syllables retained, an observation already made by Kubozono (1990, p. 12). The greater importance of SW₂ over SW₁ has also been validated by Arndt-Lappe and Plag (2013), whose findings suggest that the preservation of the prosodic structure and length of SW₂ are essential factors in determining the prosodic structure in the blend and the location of the switch point between the SWs. Because of the greater similarity between a lexical blend and its SW₂ in terms of length and stress (cf. SW₂ *lunch* and *brunch* vs. **breakfunch*, Gries 2004), we may hypothesise that SW₂ is commonly easier to access than SW₁.

However, as recently observed by Bauer (2012, p. 13), “It is generally accepted in the psycholinguistic literature that recognizability is easier for word beginnings than for word ends”. Thus, given the prototypical AD structure of blends,¹ we would expect that SW₁, whose beginning is preserved, is easier to access than SW₂, whose end is preserved instead. This expectation will not be confirmed by the results of our experiment (see, e.g., *racino* or *pleather* in § 3.3.1) since there are other factors intervening in SWs recognisability,² highly depending on how much phonological material of the SW is preserved (i.e., a syllable or less than a syllable).

In this paper, we investigate the factors that may contribute to SWs recognisability in English blends by native Italian speakers. We hypothesise that three main factors can influence the recognisability of the SWs in a blend, and therefore their understanding:

- Factor 1: Overlap at the switch point between SW₁ and SW₂.
- Factor 2: Semantic relationship – i.e., synonymy, near-synonymy, co-hyponymy, or antonymy – between SW₁ and SW₂.
- Factor 3: Contextual information referring to either SW₁ or SW₂, or both.

Experiments conducted on native speakers of Italian learning English as a Foreign Language (EFL) will be carried out to verify the validity of these

¹ See the blending formula $AB + CD \rightarrow AD$ (Plag 2003, p. 123).

² Other factors influencing SWs recognisability, such as high vs. low frequency of the SWs, their prototypicality, length, ordering, etc., will not be taken into account here.

three hypotheses. In particular, according to the first hypothesis, overlap blends, i.e., displaying overlapping segments, would be easier to access than substitution blends, where the SWs do not share any segments. According to the second hypothesis, coordinate (paradigmatic) blends, i.e., coordinating words from the same semantic field, would be more accessible than attributive (syntagmatic) blends. Finally, according to the third hypothesis, blend recognisability is facilitated in context, where contextual material anaphorically or cataphorically refers to one of the SWs or to both. A general consensus on the recognisability of the SWs in some types of English blends can 1) help distinguish the core from the periphery (Bauer 2012) and identify preferences in blend perception, and 2) understand some of the processes, such as decomposition and textual reference, which are involved in blend accessibility and interpretation.

2. Blending in English

Blending has for long been regarded as an irregular and unpredictable process in English word-formation (Aronoff 1976; Bauer 1983; Plag 2003). Despite the fact that blends are a very productive source of lexical innovation (Gries 2004, p. 639), linguists have denied them a place in regular morphology (e.g., Dressler 2000), confining them to extra-grammatical word-creation (Ronneberger-Sibold 2010). On the other hand, recent studies on English blends have shown that they display regularities in terms of prosodic structure (Bat-El, Cohen 2012; Arndt-Lappe, Plag 2013), semantic properties (Beliaeva 2014; Renner 2006), and prototypical features (Bauer 2012; Mattiello 2013, 2021). However, blends represent an under-researched area, still posing problems of classification and fuzzy boundaries (Bauer 2012).

2.1. The classification of blends

The wide variety of types and their huge diversity make blends a heterogeneous morphological category. Although several labels are used to classify them, two parameters according to which they can be categorised include their form and their meaning.

From a formal viewpoint, blends are commonly divided into two general types: **substitution** blends and **overlap** blends (Bauer 1983; Bauer, Huddleston 2002; Gries 2004; Kemmer 2003). In substitution blends, a part of one of the SWs (also called “splinter”) can be replaced with another lexeme (Lehrer 1996). For instance, *feminazi* [1989] ‘a person (typically a woman) regarded as holding extreme feminist views’ is a blend of *femi*(nist)

+ *Nazi*.³ In order to form this blend, the second part of *feminist* is replaced with another lexeme, i.e., *Nazi*. Thornton (1993, p. 145) calls this type “partial blend”, because only one of the SWs is abbreviated. Another type of substitution blend is illustrated by the blending of the first part of one SW and the final part of another. *Chugging* [2003] ← *ch(arity)* + *(m)ugging* and *racino* [1995] ← *r(acetrack)* + *(c)asino* are two recent examples of this type of blending. By contrast, overlap blending occurs when the two lexemes which are combined share a common morphological or phonological unit. For example, the combination of *sex* and *text* to make *sext* [2001], which denotes ‘a sexually explicit or suggestive message text sent electronically’, is facilitated by the /eks/ phonological segment shared by the SWs. Thus, overlap in blends helps maximisation of segments from the SWs that are preserved in the final blend.

From a semantic viewpoint, blends are divided into **coordinate** (or portmanteau) and **attributive**. For these types, Dressler (2000, p. 5) respectively uses the labels “paradigmatic contaminations” and “syntagmatic shortenings”, although he only includes the former under the heading of ‘blend’. Similarly, Plag (2003, p. 123) considers “proper blends” only the coordinate type. For instance, a *magalogue* [1978], denoting ‘a promotional catalogue designed to resemble a high-quality magazine’, coordinates *maga(zine)* and *(cata)logue*, while *beefalo* [1974], referring to ‘a cross-bred livestock animal that is three-eighths bison and five-eighths domestic cow’, is made up of *beef* and *(buff)alo*. Coordinate blends are sometimes termed “exocentric” (Bat-El 2006, p. 67) because they may not be headed. Another possible interpretation of coordinate blends is that they exhibit two heads, i.e., a *magalogue* is both ‘a type of catalogue’ and ‘a type of magazine’. Attributive or determinative blends, by contrast, are headed and, therefore, they are said to be “endocentric” (Bat-El 2006, p. 67). Like endocentric compounds, endocentric or attributive blends modify one element by another, as in *picon* [1990] ← *p(icture)* + *icon* or *glam-ma* [2003] ← *glam(our)* + *(grand)ma*. Needless to say, these two classifications, based on different parameters, can intersect, in that, morphologically, *picon* is a substitution blend and *beefalo* is an overlap blend. In *glam-ma*, the overlap is not central, but involves the SWs’ beginning (*glamour* + *grandma*).

2.2. Blending vis-à-vis compounding

While there are key differences between blends and compounds, the foundations which underpin their formation are essentially the same: the combination of established source lexemes in order to encode a new meaning.

³ Dates in square brackets and meanings of the new blends are drawn from the OED.

However, the fact that blends combine parts of lexemes, rather than whole lexemes (Kemmer 2003, p. 75), can potentially lead to problems of recognition and interpretation. While a speaker can easily identify the source lexemes of a regular compound because they are present in their entirety, this is not always the case for blends. Thus, for instance, the source lexemes of *vodka martini* are easily accessible, while the SWs of the blend *vodkatini* [1955] are less easily so.

This difference leads to the tricky issue of distinguishing blends from the neighbouring category of clipped compounds, in that the two share some type of shortening process. According to Beliaeva (2014, p. 29), clipped compounds (or “clipping compounds” as she calls them) are contractions of existing compounds (e.g., *sitcom* ← *sit(uation) com(edy)*), whereas blends are instances of creative word-formation involving the formation of new notions in the process of conceptual integration (e.g., *Oxbridge* ← *Ox(ford) + (Cam)bridge*). Thus, while in clipped compounds shortening takes place after compounding, in blending shortening and compounding happen simultaneously. Other scholars rather discriminate between the two categories on the basis of their different structure: i.e., blends are generally AD-forms, whereas clipped compounds commonly conform to an AC pattern (Bat-El 2006).

The fact that the two source lexemes of a clipped compound are often attested together (e.g., *situation comedy*) whereas the SWs of a blend are not (e.g., **Oxford Cambridge* is not a compound) increases the importance of recognisability in the latter. Morphologists indeed agree that one of the fundamental notions in blends is their recognisability: i.e., “[b]oth elements in the blend must be recognizable if the blend is to be successful” (Bauer 2012, p. 13). The accessibility and acceptability of novel blends have been already tested by Lehrer (1996) and Connolly (2013) with native speakers. However, the degree of SWs recognisability by non-native speakers has not been investigated hitherto, nor have blends been classified and distinguished on the basis of their optimal structure for perception and interpretation. The experiment conducted in this paper is meant to fill this gap.

3. The recognisability of blend source words/meanings: An experiment

An experiment was carried out in order to investigate the ways in which non-native speakers identify the source lexemes of English blends and how increased recognisability can favour the accessibility of some types of blends rather than others. In the experiment, the participants were presented with a number of blends and were asked to identify their source lexemes, either SW₁

or SW₂, with no time constraints. The following subsections outline key aspects of the experiment.

3.1. Participants

18 Italian students of EFL participated in the experiment. The participants were all native speakers of Italian from various parts of Italy, and all were between 20-23 years old. Of these participants, 10 were male and 8 were female. All participants were attending University in Pisa and had not completed their first cycle of university education. They all had a certified B2 CEFR level.

3.2. Materials and methods

The materials used for the experiment included a list of 36 blends selected from a larger database of recent instances. All blends were existing words attested in the OED and had been obtained via an advanced search on the dictionary platform.⁴ Recent blends were chosen because they are not widely known, especially by Italian speakers, who are generally not much familiar with the blending process (Thornton 2004). However, results from participants who were familiar with the blends presented were excluded from the analysis.

The experiment was divided into three parts, with a ten-minute break between them. In each part, the participants were presented with a list of twelve blends. In the first two parts, the blends were accompanied by either their SW₁ or their SW₂. In other words, one of the SWs was given in order to activate the process of recognition of the other SW, either by subtracting the disclosed SW from the total blend, or by associating the disclosed SW's meaning with the meaning of the undisclosed SW. In the third part, the blends were given in a context from COCA (*Corpus of Contemporary American English*) or NOW (*News on the Web*) corpora. The stimuli were presented to the subjects in random order with or without context, in an attempt to access the ways in which the participants assessed the blends' meanings based solely on their structure or aided by co-textual information.

3.3. Procedure

The experiment described in this paper aimed to investigate the interpretative strategies of the participants. The aim was not of psycholinguistic nature, but to examine how diverse types of blends can have implications for their

⁴ The filter 'blend' was used to select all the entries which are labelled blends in the OED's etymology. Only recent blends attested from 1950 onwards were chosen for the experiment.

interpretations. This article explores whether or not the recognisability of the source lexemes of a blend and its co-textual material might be factors influencing the blend accessibility, and, more generally, suggest preferences or trends in blend perception and recognition.

In particular, the experiment aimed to investigate a number of hypotheses, which are outlined as follows:

- if the blend displays an overlap at the switch point between the SWs, a majority of speakers will agree on its source lexemes and on its meaning;
- where the SWs in a blend are in a semantic relationship of identity (e.g., fantastic and fabulous), similarity (e.g., jeans and leggings), or opposition (e.g., friend vs. enemy), the more likely it is that a consensus will be achieved with regard to its source lexemes and meaning. In the above cases, tautology is expected to be the easiest to recognise while antonymy to be the most difficult;
- where the SWs in a blend are in an anaphoric or cataphoric reference relation with the co-textual elements of the blend, the more likely it is that a consensus will be achieved with regard to its source lexemes and meaning.

Each participant was given a questionnaire divided into three parts. When each part was concluded, it was not possible to change any of the previous responses. The questionnaire gave a definition of a blend and cited two examples of the phenomenon. Participants were presented with one of the three lists, the first two lists with twelve blends, the last one with eleven blends. For the first two parts, six of the blends on each list were followed by their SW₁ and the other six by their SW₂. Participants were asked to cite the missing source lexemes and encouraged to guess the blends' meanings. The meanings were explained in Italian by the participants and then translated into English by the present author. Only the most cited meanings were taken into account. For the last part, all the blends were followed by a contextualised example but no SWs. Here participants were asked to guess the blends' meanings based on co-textual information.

The blends were of a number of different types. In the first list, six were formed through substitution and six were formed through overlap. Blends with a non-central overlap or only graphic overlap were considered substitution blends.⁵ Of overlap blends, two were partial blends preserving one SW in its entirety (indicated with an asterisk in Table 1) and one preserved both SWs (indicated with two asterisks). In both the second and the

⁵ According to Beliaeva (2014, p. 59), a non-central overlap occurs when the SWs have one or more coinciding letters/phonemes either at the beginning (*snarfle* ← *snarf* + *snaffle*) or at the end (e.g., *hoolivan* ← *hooligan* + *van*).

third list, six attributive blends and six (Table 2)/five (Table 3) coordinate blends were presented. The number of syllables in the blends in each list also varied from one to four syllables. The blends in Tables 2-4 are given in order of length, from monosyllabic to four-syllable blends. Table 1 shows the blends which were used in the experiment and the various categories they represent. They are listed in alphabetical order.

Types	Overlap blends	Substitution blends
Attributive blends	bromance** Clintonomics* dancercise flexecutive* flexitarian freegan* freemium* glamping militician monergy sext* shoppertainment* vog webisode* webliography*	gengineer* glam-ma hoolivan* machinima wigger
Coordinate blends	animatic beefalo* boatel* burkini frenemy* glocal* hip-hopera** jeggings magalogue pleather*	fantabulous racino skort smaze zonkey

Table 1

The blends included in the experiment. Those marked with * preserve one of their SWs entirely and those marked with ** preserve both SWs with an overlap.

3.3.1. Hypothesis 1: Overlap between the source lexemes entails consensus about the source words/meaning

In order to examine hypothesis 1, the participants were presented six overlap blends and six substitution blends in random order. The first SW was disclosed for half of the blends and the second SW for the other half, sometimes facilitating recognition, other times making it harder. Then, participants were asked to provide the missing SWs and the meanings of the blends. All participants declared that they were not familiar with the blends. The experiment indeed seeks to investigate how non-native speakers interpret innovative blends; therefore, unfamiliarity was an essential prerequisite for relevance to this study.

Table 2 shows the participants' responses to the question about the source lexemes/meanings of the blends. For the undisclosed SWs, each response given is followed by the number of respondents who provided that

response. The correct response (provided by the OED) is underlined in the table. Some of the slots were left incomplete by the participants. For the blends' sense, only the most commonly-cited meanings have been highlighted.

Blend	Source Word 1	Source Word 2	Most commonly-cited meaning
smaze	<i>small</i> x 8 <i>smack</i> x 3 <i>smart</i> x 3 <u><i>smoke</i></u> x 2 (2 incomplete)	haze	not much haze x 7 haze mixed with smoke x 2
bromance	bro	<u><i>romance</i></u> x 16 <i>performance</i> x 2	a romance between brothers x 14 the performance of one's bro x 1
freegan	free	<i>slogan</i> x 10 <u><i>vegan</i></u> x 4 <i>hooligan</i> x 2 <i>organ</i> x 1 (1 incomplete)	a free slogan x 9 a person who generally eats no animal products x 4
pleather	No conclusive response (7 different responses, 11 incomplete) <u><i>plastic</i></u>	leather	No meanings provided
glam-ma	glamorous	<i>cinema</i> x 4 <u><i>grandma</i></u> x 1 <i>panorama</i> x 1 <i>schema</i> x 1 (11 incomplete)	a glamorous cinema x 4 a glamorous grandmother x 1
wigger	<i>woman</i> x 13 <i>wild</i> x 2 (3 incomplete) <u><i>white</i></u>	nigger	a black woman x 12
monergy	<i>monitor</i> x 6 <u><i>money</i></u> x 5 <i>month</i> x 1 (6 incomplete)	energy	a monitor powered by energy x 4 energy/power of money x 3
magalogue	magazine	<u><i>catalogue</i></u> x 10 <i>dialogue</i> x 8	a catalogue resembling a magazine x 8
hoolivan	<u><i>hooligan</i></u> x 17 (1 incomplete)	van	a van for hooligans x 15
racino	<i>race</i> x 7 <i>racism</i> x 2 <i>racist</i> x 1 <i>racial</i> x 1 (7 incomplete) <u><i>racetrack</i></u>	casino	a casino where people race x 4
machinima	machine	<i>minima</i> x 2 <i>anima</i> x 2 (14 incomplete) <u><i>cinema</i></u>	No meanings provided
Clintonomics	Clinton	<u><i>economics</i></u> x 18	economics of President Clinton x 17

Table 2

Identification of the source lexemes of overlap vs. substitution blends. The disclosed SW is marked in bold, the undisclosed SW(s) provided by respondents is/are in italics, the correct one is also underlined.

The most common responses to the question about source lexemes and meanings allowed us to verify hypothesis 1: if a blend displays an overlap at the switch point, a majority of participants will agree on its source lexemes/meaning. For SW₁ in **overlap blends**, 28% of respondents provided *money* for *monergy*, but no response *plastic* was provided for *pleather*. For SW₂, 100% of respondents provided *economics* for *Clintonomics*, 89% provided *romance* for *bromance*, 56% provided *catalogue* for *magalogue*, and 22% suggested *vegan* for *freegan*. Even if with a low number of participants percentages below 80% have little or no significance, it is worth noting that SW₂ was recognised by the participants in the majority of cases. This may be because it generally provides the stress and prosodic contour for the final blend. Moreover, thanks to the overlap, the identification of the correct SW also facilitated meaning recognition. The overlapping phonemes ranged from one in *free* + *vegan* /i:/, *plastic* + *leather* /l/ and *money* + *energy* /n/ to two in *bro* + *romance* /rəʊ/ and *Clinton* + *economics* /ən/. In *magazine* + *catalogue* /æ ə/, a discontinuous overlap facilitated SW recognition. From these findings we can infer that, not only the overlap, but also its position in the SWs is fundamental in the process of SW recognition: indeed, in *pleather*, the overlap occurs between the onsets of the first syllables of *plastic* and *leather*. Hence, while the latter is entirely identifiable, the former is scarcely recognisable. By contrast, in *bromance*, the overlap allows for the preservation of both SWs because the switch point is between the only syllable of SW₁ and the first syllable of SW₂.

In overlap blends, the percentages of respondents giving a (nearly) correct meaning progressively were: *Clintonomics* ‘the economic policies of President Clinton’ (94%), *bromance* ‘intimate friendship between men’ (78%), *magalogue* ‘promotional catalogue designed to resemble a magazine’ (44%), and *freegan* ‘a person who eats discarded food’ (22%). For *monergy* ‘expenditure on energy’, the meanings suggested by the participants were far from the correct one, while for *pleather* ‘a synthetic fabric treated to resemble leather’ no meaning was provided because the respondents were not able to identify the undisclosed SW₁ *plastic*. It is not surprising that there is a strict correlation between SWs recognisability and meaning understanding.

For the SWs in **substitution blends**, respondents displayed many more solutions than for overlap blends. For SW₁, 94% of respondents provided *hooligan* for *hoolivan*, mainly because *hooli-* is an infrequent word beginning. However, there was much more uncertainty for the other SWs: e.g., only 11% provided *smoke* for *smaze*, yet most respondents suggested *small*, and a considerable number *smack* and *smart*, probably because they assumed that a graphical overlapping *-a-* might link the two SWs. A similar uncertainty was for SW₁ in *racino*, with four concurrent responses, i.e., *race* (39%), *racism* (11%), *racist* (6%), *racial* (6%), and 39% left incomplete, the

correct one being *racetrack*. For *wigger*, no respondents suggested correct *white*, but 72% offered *woman* and 11% proposed *wild*, the latter response perhaps partially guided by a graphical overlapping *-i-*. For SW₂, only 6% suggested correct *grandma*, while *cinema* (22%), *panorama* (6%), and *schema* (6%) were alternative options, 61% left the slot empty. Similarly, for *machinima*, 78% left the response incomplete and correct *cinema* was not identified, *minima* and *anima* rivalling with 11% each. In this case, the blend respelling *machinima* (in the place of *machinema*) may have obstructed the recognition of SW₂.

The difficulty in SW recognition correlated with difficulty in meaning identification: *smaze* ‘a mixture of smoke and haze’ (11%), *glam-ma* ‘a glamorous grandmother’ (6%), no meaning for *machinima*, and incorrect meaning (based on incorrect SWs) for *wigger* and *hoolivan*. For *racino*, the correct SW (*racetrack*) was not identified by the respondents, but the majority of cases provided the more general term *race*, which allowed them to draw near the meaning ‘a building complex having a racetrack and gambling facilities associated with casinos’. These percentages are not comparable to those recorded for overlap blends.

Based on these data, it appears that we can draw some inferences about the effect that overlap has both on consensus on the source lexemes of a blend and on related consensus about its meaning. There appears that consensus on the correct source lexemes is most commonly reached where the blend SWs share from one to three phonemes, also discontinuously. Even if further investigation with larger datasets would be necessary before drawing firmer conclusions, we can also infer that consensus on SWs correlates with consensus about the blend’s meaning.

3.3.2. Hypothesis 2: Semantic relationship of similarity (or opposition) between the blend SWs entails consensus about the source words/meaning

In order to examine hypothesis 2, the participants were presented six coordinate blends and six attributive blends in random order. The first SW was disclosed for half of the blends and the second SW for the other half, sometimes facilitating recognition, other times making it harder. Then, participants were asked to provide the missing SWs and the meanings of the blends. All participants declared that they were not familiar with the blends, except for *webliography*, which was known by three respondents, and *fantabulous*, known by one respondent.⁶ Table 3 shows the participants’ responses to the question about the source lexemes/meanings of the blends.

⁶ These specific results have been excluded from Table 3.

For the undisclosed SWs, each response given is followed by the number of respondents who provided that response. As above, the correct response is underlined in the table. Some of the slots were left incomplete by the participants. For the blends' meanings, only the most commonly-cited meanings have been highlighted.

Blend	Source Word 1	Source Word 2	Most commonly-cited meaning
vog	<i>vague</i> x 3 (15 incomplete) <i>volcanic</i>	fog	No meanings provided
skort	<i>skirt</i> x 16 <i>skate</i> x 2	short(s)	shorts resembling a skirt x 7 both skirt and shorts x 8
glamping	glamorous	<i>camping</i> x 10 <i>jumping</i> x 3 (5 incomplete)	glamorous, fashionable camping x 9 high jump x 2
zonkey	<i>zebra</i> x 15 <i>zone</i> x 2 (1 incomplete)	donkey	the offspring of a zebra and a donkey x 10 an animal resembling a zebra and a donkey x 4
jeggings	jeans	<i>leggings</i> x 18	leggings that resemble a pair of jeans x 16 tight-fitting, skinny jeans x 2
frenemy	<i>friend</i> x 15 <i>friendly</i> x 3	enemy	a person who is both a friend and an enemy x 15 a false friend x 3
gengineer	<i>general</i> x 14 (4 incomplete) <i>genetic</i>	engineer	a general engineer x 6
hip-hopera	hip-hop	<i>opera</i> x 18	music that combines hip-hop and opera x 16 opera with hip-hop music x 2
flexitarian	flexible	<i>humanitarian</i> x 4 <i>authoritarian</i> x 4 (10 incomplete) <i>vegetarian</i>	a person who is authoritarian, but also flexible x 2
militician	military	<i>technician</i> x 4 <i>politician</i> x 2 <i>musician</i> x 2 (10 incomplete)	technician involved in the military x 1 politician involved in the military x 1
fantabulous	<i>fantastic</i> x 17 (1 excluded)	fabulous	fantastic and fabulous x 17
webliography	web	<i>bibliography</i> x 8 (4 incomplete, 6 excluded)	bibliography of web sources x 6 bibliography on the web x 2

Table 3

Identification of the source lexemes of coordinate vs. attributive blends. The disclosed SW is marked in bold, the undisclosed SW(s) provided by respondents is/are in italics, the correct one is also underlined.

The most common responses to the question about source lexemes and meanings allowed us to verify hypothesis 2: if the SWs of a blend display semantic similarity, a majority of participants will agree on its source lexemes/meaning. For SW₁ in **coordinate blends**, 94% of respondents provided *fantastic* for *fantabulous*, 89% provided *skirt* for *skort*, 83% provided *zebra* for *zonkey*, and 83% suggested *friend* for *frenemy*. Although the phonological similarity between SW₁ and the blends *fantabulous* and *skort* may respectively have helped the recognition of *fantastic* and *skirt*, the synonymy relationship with disclosed *fabulous* and the co-hyponymy relationship with *shorts* have undoubtedly contributed to the SWs identification. The importance of a semantic relationship between the SWs in coordinate blends is even more evident in *zonkey*, in which *zebra* has been mainly identified thanks to its co-hyponymy with *donkey*, and in *frenemy*, in which the antonymic relationship between *friend* and *enemy* has facilitated the respondents' task. This contradicts expectations that antonymy is the most difficult relation to recognise (cf. Section 3.3). In addition, a semantic relationship of co-hyponymy has helped participants in the recognition of SW₂ in the coordinate blends *jeggings* and *hip-hopera*, whose second words *leggings* and *opera* were identified by all respondents (100%).

In coordinate blends, the meaning identification task also appeared facilitated by the SWs' semantics. The percentages of respondents giving a (nearly) correct meaning progressively were: *fantabulous* 'of almost incredible excellence' (100%), *jeggings* 'tight-fitting stretch leggings styled to resemble a pair of denim jeans' (89%), *hip-hopera* 'music that combines elements of hip-hop and opera' (89%), *frenemy* 'a person with whom one is friendly, despite a fundamental rivalry' (83%), *skort* 'a pair of shorts having a flap across the front to give the appearance of a skirt' (83%), and *zonkey* 'the offspring of a zebra and a donkey' (55%).

The same easiness was not evidenced by the results for **attributive blends**. Recognition of SW₁ was highly difficult in *vog*, in which nobody suggested the correct word *volcanic*, and in *gengineer*, for which most suggested *general* (78%) instead of *genetic*. Recognition of SW₂ was difficult in *militician*, for which only 11% provided *politician*,⁷ and impossible in *flexitarian*, for which nobody suggested *vegetarian*,⁸ easier for *glamping* (*camping* was provided by 55%) and *webliography* (with *bibliography* given by 66% of the participants who did not know the blend). The latter two blends were assigned the correct meaning by 50% (*glamping*) and 66%

⁷ Although this is considered an attributive blend, the two SWs are semantically/pragmatically related, as in most dictatorships the rulers (or politicians) are military men.

⁸ The fact that 4 respondents provided *authoritarian* as SW₂ means that language users/learners strive to establish semantic connections between SWs (in this case an antonymic one) to make sense of novel blends. This piece of evidence bears on hypothesis 2.

(*webliography*), as a further confirmation that SWs recognition helps meaning identification.

Based on these results, it can be said that also hypothesis 2 has been confirmed. While it appears that participants have most difficulty in coming to a consensus about the source lexemes and meaning of attributive blends, it also appears that the degree of consensus increases with coordinate blends. The degree of consensus with regard to the SWs of coordinate blends overcomes 83% for all blends, and their meaning has been identified by more than 50% of the participants, for five of them by more than 80%. In attributive blends, the ambiguity of competition between possible SWs increases compared to coordinate blends, whose SWs are more similar in syntactic and semantic terms.

3.3.3. Hypothesis 3: Co-textual information (e.g., reference to the blend SWs) entails consensus about the blend's meaning

In order to examine hypothesis 3, the participants were presented five coordinate blends and six attributive blends randomly, first in isolation and then in context. The blend SWs were not disclosed, but, on some occasions, they were mentioned in the contexts provided. In one context, the SWs appeared entirely and in proximity to (*esp. following*) the blend.

Then, participants were asked to provide the meanings of the blends. All participants declared that they were not familiar with the blends. Table 3 shows the participants' responses to the question about the meanings of the blends. The blend SWs are reported in the table, but, as said, they were not provided to the respondents. Each response given in the last column is followed by the number of respondents who provided that response.

Blend	Context	Most commonly-cited meaning
sext <i>sex + text</i>	Whether a sext qualifies as relatively safe sexual experimentation or a disaster often depends on who finds out about it. (COCA, 2014)	a sexual text message x 13 a text about sex x 3 (2 incomplete)
glocal <i>global + local</i>	The inventive expression, ' glocal ', speaks to the intersections of the global and the local, how predominantly local events are globalised, and those of global resonances are reinterpreted or reformulated in local, even rustic settings. (NOW, 2019)	both global and local x 17 in-between global and local x 1
boatel <i>boat + hotel</i>	Now, Ireland is set to get its own boatel – and it could be here as early as the end of this year. (NOW, 2017)	a small boat x 3 (15 incomplete)

freemium <i>free + premium</i>	Amazon is reportedly developing a freemium version of Prime Video. (COCA, 2017)	something free x 8 (10 incomplete)
dancercise <i>dance + exercise</i>	For all those who are tired of going to the gym and getting on the boring treadmill, the dancing diva will be sharing her secrets of staying in shape through a special ' Dancercise ' workout. (NOW, 2016)	dance and exercise x 10 a dancing exercise x 5 exercise while dancing x 3
beefalo <i>beef + buffalo</i>	The Fort served perfectly medium-rare beefalo . (COCA, 2004)	beef x 5 meat such as beef x 2 (11 incomplete)
burkini <i>burka + bikini</i>	She was also the first woman to wear a burkini during the swimsuit portion of the event. (COCA, 2017)	bikini in the form of a burka x 7 swimsuit that resembles a burka x 3 (8 incomplete)
webisode <i>web + episode</i>	Nimbus became one of the early experimenters with the short online film, or " webisode ". (COCA, 2012)	short online film x 16 episode on the web x 2
animatic <i>animated + schematic</i>	We've reached out to Fox for an official comment on the animatic , and will update this post if we hear a response. (COCA, 2016)	animated cartoon x 4 (14 incomplete)
flexecutive <i>flexible + executive</i>	<ul style="list-style-type: none"> • Realities of the Personal Information Economy • Birth of the new Flexecutive • Development of new Bleisure Hives (NOW, 2014) 	executive x 7 (11 incomplete)
shoppertainment <i>shopper + entertainment</i>	This is unlike visits to larger malls, which are for ' shoppertainment ', not necessarily to make a purchase. (COCA, 2002)	shopping mall x 3 entertainment for shoppers x 1 (14 incomplete)

Table 4

Identification of the meaning of coordinate vs. attributive blends in context.

The most common responses to the question about the blends' meanings allowed us to verify hypothesis 3: if the context of a blend provides references to its SWs, a majority of participants will agree on its meaning. The results reported in Table 4 point out that all contexts that provided no relevant information about the blend SWs were unhelpful for the participants. For instance, for *boatel* 'boat which functions as a hotel', 83% left the task incomplete, and the remaining respondents only identified SW₁ (*boat*), as evidenced by the meaning suggested. This result is surprising, because SW₂ (*hotel*) is an Anglicism that Italians know and regularly use; therefore, we would have expected it to be identified by at least some of the respondents.

Similarly, for *freemium* 'a free gift, given by a business in order to persuade customers to pay for other goods or services', 55% left the task

incomplete and 44% only recognised SW₁ (*free*), while for *animatic* ‘a preliminary mock-up of a film’, 78% did not provide any meaning, and the remaining 22% included in the meaning suggested SW₁ (*animated cartoon*), but not SW₂ (*schematic*). By contrast, for *flexexecutive* ‘a professional whose use of information technology offers flexible employment opportunities’, only SW₂ (*executive*) was identified by 39%, but not the first SW, probably due to the excessive brevity of the context.

On the other hand, when the two SWs were cited in the context (*global* and *local*), all respondents (100%) identified the meanings of *glocal* ‘both global and local’. Other words provided by the contexts functioned as semantic indices for disambiguation. In *burkini* ‘swimsuit for women which covers the head and body’, *swimsuit* was an anaphoric semantic prime for *bikini*, identified by 40% of respondents, while *short online film* was a cataphoric semantic prime for *webisode*, whose meaning ‘an episode of a drama or comedy series, which is made available online’ was understood by all respondents (100%), although only 11% recognised the two SWs *web* and *episode*. Partial anaphoric reference (*sexual*) helped the comprehension of *sext* ‘a sexually explicit message sent electronically’ (89%), while the verb *served* was less useful for the understanding of *beefalo* ‘a cross-bred livestock animal that is three-eighths bison and five-eighths domestic cow’, whose SW₁ *beef* was clearer to respondents (39%) than SW₂ *buffalo* (0%).

For *shoppertainment* ‘the provision of entertainment within a shopping centre’, *malls* and *purchase* may even have acted as distractors for those who answered ‘shopping mall’ (16%), while only 5% identified SW₂ *entertainment*. Lastly, for *dancercise* ‘dancing performed as an exercise’, a combination of anaphoric and cataphoric semantic primes (*gym*, *dancing*, *staying in shape*, *workout*) helped the identification of both SWs *dance* and *exercise* by 100% of respondents. Needless to say, an anaphoric relation (as with *global* and *local* referring back to *glocal*) is easier to identify than a cataphoric relation (as with *dancing* preceding *dancercise*). Moreover, the adjacency (vs. distance) of the blends to their SWs (or their semantic primes) may have also played a role in the recognition and interpretation processes.

4. General discussion

A comparison of the results obtained from the three-part experiment gives a clearer idea of the factors influencing blends’ interpretation. Table 5 compares the percentages for the source lexemes and correct meanings of overlap vs. substitution blends, coordinate vs. attributive blends, and blends found in informative vs. uninformative context. When the context was found to be a distractor, it was considered uninformative.

Blend	Comparison of % of correct source lexemes responses	Comparison of % of correct meaning responses
Overlap vs. substitution	49% vs. 18%	43% vs. 6%
Difference between overlap vs. substitution	D = 31%	D = 37%
Coordinate vs. attributive	92% vs. 18%	75% vs. 15%
Difference between coordinate vs. attributive	D = 74%	D = 60%
Informative context vs. Uninformative context	74% vs. 1%	88% vs. 1%
Difference between informative vs. uninformative context	D = 73%	D = 87%

Table 5

Percentages of correct source lexemes/meaning responses in the three-part experiment.

These results show that consensus about source lexemes and meaning is more readily reached when 1) the blend exhibits an overlap between the SWs, 2) the two SWs are semantically related, and 3) the blend is found in a relevant context providing semantic prime(s) to either its SWs or its meaning. In particular, the results of a comparison between the different types of blends that functioned as stimuli in the experiment show that prior knowledge of one of the SWs of a coordinate blend can significantly help (92%) in the identification of the other SW (e.g., *shorts* → *skirt*, *donkey* → *zebra*), while in an attributive blend the two SWs are much more independent and harder to identify, because they are semantically unrelated (cf. *volcanic* vs. *fog*, *genetic* vs. *engineer*), only syntactically related by a syntagmatic grammatical relation.

When the SWs are provided by the context or semantic primes are given to respondents in the blend's co-text, their recognition is high (74%) and the meaning readily identified (88%), whereas a concise and uninformative context is irrelevant to blend recognisability.

Finally, the presence of an overlap at the switch point has revealed to be helpful for both SWs identification (49%) and meaning disambiguation (43%). Indeed, an overlap allows maximisation of segments from the SWs that are preserved in the final blend (e.g., *Clinton* + *economics*). However, when the overlap occurs very early in SW₁ (*plastic* + *leather*) or very late in SW₂ (*boat* + *hotel*), maximisation of segments is jeopardised by optimal structure, and only one of the SWs is recognised (*leather*, *boat*).

5. Conclusions

The hypotheses which were proposed before the experiment generally appear to have been borne out by the results. The results show a direct correlation between consensus on the identification of source lexemes and on the identification of the meaning for blends with an overlap between the SWs. There is even more evidence to support the assertion that consensus on the SWs/meaning is most easily reached with coordinate blends whose SWs stand in a co-hyponymy relationship, or with blends whose SWs are cited in the co-text, either as anaphoric or as cataphoric references.

It must be noted that this is a pilot experiment with a small number of participants and a small number of blends and larger numbers of participants and/or blends might have led to different results. The main aim of this experiment was to explore the nature of consensus in assigning meaning to innovative English blends by non-native speakers learning EFL. In general, the experiment confirmed that blends are hard to disambiguate for Italian learners of EFL and that their SWs are often indiscernible because of the economical nature of blends compared to compounds. However, the context where a blend is used can significantly help in blend recognisability and learners should check for primes before or after the blend in order to find out formal or semantic similarities between the blend itself and its co-textual information.

Ultimately, the experiment can allow us to draw some conclusions about the distinction between the core and the periphery in blend perception and recognition. Semantically, coordinate blends, whose SWs are paradigmatically rather than syntagmatically related, are easier to perceive and recognise than attributive blends (cf. Mattiello 2021 for the preferred semantic types of blends in the OED). Formally, overlap blends mixing SWs that share some segments and merging them where they overlap, are easier to recognise than substitution blends. Although generalisations would require statistically more significant figures, we can infer from our results that overlap and coordinate blends belong to the core for perception. Contextually, the accessibility of blends may be favoured by the co-presence of at least one of the source lexemes, or of synonyms which may act as semantic primes to the source lexemes.

In general, the meaning of an unknown form can often be reconstructed in the context of a broader discussion. However, blend recognisability and accessibility seem to require specific reference to their source lexemes, in that other related words may even function as distractors to the specific blend's meaning. Other variables can of course influence blend recognisability, such as the frequency and length of the source lexemes, rhyme between SW₂ and the blend, and other structural aspects concerning the combination of the source lexemes. Further experiments can take these variables into account,

but it appears that the context (or co-text) remains the main factor influencing blend accessibility.

Bionote: Elisa Mattiello is Associate Professor of English Language and Linguistics at the University of Pisa. She holds a Ph.D. in English Linguistics from the same University, where she teaches undergraduate courses of ESP. Her research focuses on English word-formation and lexicology, with particular attention to the creative mechanisms that are used to coin new words in English. She has published in leading journals, including *Linguistics*, *Lingua*, *Languages in Contrast*, *Language & Communication*, and *International Journal of Language Studies*. She authored the monographs on slang (Polimetrica 2008), *Extra-grammatical Morphology in English* (De Gruyter, 2013), *Analogy in Word-formation* (De Gruyter, 2017), *Linguistic Innovation in the Covid-19 Pandemic* (Cambridge Scholars Publishing, 2022), and *Transitional Morphology* (Cambridge University Press, 2022). She is currently part of the Editorial Boards of the *International Journal of English Linguistics*, *International Journal of Language and Linguistics*, and *Cambridge Scholars Publishing*.

Author's address: elisa.mattiello@unipi.it

References

- Arndt-Lappe S. and Plag I. 2013, *The role of prosodic structure in the formation of English blends*, in “English Language and Linguistics” 17, pp. 537-563.
- Aronoff M. 1976, *Word Formation in Generative Grammar*, MIT Press, Cambridge, MA.
- Bat-El O. 2006, *Blends*, in Brown K. (ed.), *Encyclopedia of Language and Linguistics*, Elsevier, Oxford, pp. 66-70.
- Bat-El O. and Cohen E.-G. 2012, *Stress in English blends: A constraint-based approach*, in Renner V., Maniez F. and Arnaud P. (eds.), *Cross-disciplinary Perspectives on Lexical Blending*, De Gruyter, Berlin, pp. 193-211.
- Bauer L. 1983, *English Word-formation*, Cambridge University Press, Cambridge.
- Bauer L. 2012, *Blends: Core and periphery*, in Renner V., Maniez F. and Arnaud P.J.L. (eds.), *Cross-Disciplinary Perspectives on Lexical Blending*, De Gruyter, Berlin, pp. 11-22.
- Bauer L. and Huddleston R. 2002, *Lexical word-formation*, in Huddleston R. and Pullum G.K. (eds.), *The Cambridge Grammar of the English Language*, Cambridge University Press, Cambridge, pp. 1621-1723.
- Beliaeva N. 2014, *A study of English blends: From structure to meaning and back again*, in “Word Structure” 7, pp. 29-54.
- Cannon G. 1986, *Blends in English word formation*, in “Linguistics” 24, pp. 725-753.
- COCA: *Corpus of Contemporary American English* 1990-2017. <https://corpus.byu.edu/coca/> (03/04/2019).
- Connolly P. 2013, *The innovation and adoption of English lexical blends*, in “JournLIPP” 2, pp. 1-14.
- Dressler W.U. 2000, *Extragrammatical vs. marginal morphology*, in Doleschal U. and Thornton A.M. (eds.), *Extragrammatical and Marginal Morphology*, Lincom Europa, München, pp. 1-10.
- Gries S.Th. 2004, *Isn't that fantabulous? How similarity motivates intentional morphological blends in English*, in Achard M. and Kemmer S. (eds.), *Language, Culture, and Mind*, CSLI, Stanford, CA, pp. 415-428.
- Gries S.Th. 2006, *Cognitive determinants of subtractive word-formation processes: A corpus-based perspective*, in “Cognitive Linguistics” 17 [4], pp. 535-558.
- Kemmer S. 2003, *Schemas and lexical blends*, in Cuyckens H.C., Berg T., Dirven R. and Panther K.-U. (eds.), *Motivation in Language: From Case Grammar to Cognitive Linguistics. Studies in Honour of Günter Radden*, Benjamins, Amsterdam/Philadelphia, pp. 69-97.
- Kubozono H. 1990, *Phonological constraints on blending in English as a case for phonology-morphology interface*, in “Yearbook of Morphology” 3, pp. 1-20.
- Lehrer A. 1996, *Identifying and interpreting blends: An experimental approach*, in “Cognitive Linguistics” 7 [4], pp. 359-390.
- Martinet A. 1955, *Economie des changements phonétiques. Traité de phonologie diachronique*, A. Francke, Bern.
- Mattiello E. 2013, *Extra-grammatical Morphology in English. Abbreviations, Blends, Reduplicatives, and Related Phenomena*, De Gruyter, Berlin/Boston.
- Mattiello E. 2021, *Blends vis-à-vis compounds in English*, in “Italian Journal of Linguistics / Rivista di Linguistica” 33 [2], pp. 131-162.
- NOW: *News on the Web* 2010-2018. <https://corpus.byu.edu/now/> (03/04/2019).
- OED2-3: *Oxford English Dictionary Online* 1989-2019, 2nd/3rd ed. Oxford: Oxford University Press. <http://www.oed.com/> (03/04/2019).

- Plag I. 2003, *Word-formation in English*, Cambridge University Press, Cambridge.
- Renner V. 2006, *Les composés coordinatifs en anglais contemporain*, Dissertation, Université Lumière Lyon 2, Lyon.
- Ronneberger-Sibold E. 2010, *Word creation: Definition – function – typology*, in Rainer F., Dressler W.U., Kastovsky D. and Luschützky H.C. (eds.), *Variation and Change in Morphology*, Benjamins, Amsterdam/Philadelphia, pp. 201-216.
- Thornton A.M. 1993, *Italian blends*, in Tonelli L. and Dressler W.U. (eds.), *Natural Morphology. Perspectives for the Nineties*, Unipress, Padova, pp. 143-155.
- Thornton A.M. 2004, *Parole Macedonia*, in Grossman M. and Rainer F. (eds.), *La formazione delle parole in italiano*, Niemeyer, Tübingen, pp. 567-571.
- Zipf G.K. 1949, *Human Behavior and the Principle of Least Effort*, Cambridge, Addison-Wesley Press.